

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
20 September 2001 (20.09.2001)

PCT

(10) International Publication Number
WO 01/68826 A2

(51) International Patent Classification⁷: C12N 15/00

(74) Agent: BASTIAN, Werner; Syngenta Participations AG,
Intellectual Property, P.O. Box, CH-4002 Basel (CH).

(21) International Application Number: PCT/EP01/02753

(22) International Filing Date: 12 March 2001 (12.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/525,152 14 March 2000 (14.03.2000) US

(71) Applicant (*for all designated States except US*): SYN-
GENTA PARTICIPATIONS AG [CH/CH]; Schwarzwald-
allee 215, CH-4058 Basel (CH).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): FONNE-PFISTER,
Raymonde [FR/FR]; 6, rue de Héringue, F-68220 Folgès-
bourg (FR). DE MARCO, Ario [IT/DE]; Am Waldrand 10,
D-69126 Heidelberg-Boxberg (DE). VOLRATH, Sandra,
Lynn [US/US]; 4225 Pin Oak Drive, Durham, NC 27707
(US).

Published:

— without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.



WO 01/68826 A2

(54) Title: PROTOPORPHYRINOGEN OXIDASE ("PROTOX") GENES

(57) Abstract: The present invention provides novel DNA sequences coding for protoporphyrinogen oxidase (protox) enzymes from maize, soybean, wheat, cotton, sugar beet, oilseed rape, rice, sorghum, and sugar cane. In addition, the present invention teaches modified forms of protox enzymes that are herbicide tolerant. Plants expressing herbicide tolerant protox enzymes taught herein are also provided. These plants may be engineered for resistance to protox inhibitors via mutation of the native protox gene to a resistant form or they may be transformed with a gene encoding a herbicide tolerant form of a plant protox enzyme.

PROTOPORPHYRINOGEN OXIDASE ("PROTOX") GENES

The present invention relates to DNA molecules encoding protoporphyrinogen oxidase ("protox") enzymes and to DNA molecules encoding herbicide-tolerant forms of the enzyme protoporphyrinogen oxidase ("protox"). The invention further relates to herbicide-tolerant plants as well as methods for tissue culture selection and herbicide application based on these herbicide-tolerant forms of protox.

I. The Protox Enzyme and its Involvement in the Chlorophyll/Heme Biosynthetic Pathway

The biosynthetic pathways that lead to the production of chlorophyll and heme share a number of common steps. Chlorophyll is a light harvesting pigment present in all green photosynthetic organisms. Heme is a cofactor of hemoglobin, cytochromes, P450 mixed-function oxygenases, peroxidases, and catalyses (*see, e.g. Lehninger, Biochemistry*, Worth Publishers, New York (1975)), and is therefore a necessary component for all aerobic organisms.

The last common step in chlorophyll and heme biosynthesis is the oxidation of protoporphyrinogen IX to protoporphyrin IX. Protoporphyrinogen oxidase (referred to herein as "protox") is the enzyme that catalyzes this last oxidation step (Matringe *et al.*, *Biochem. J.* 260: 231 (1989)).

The protox enzyme has been purified either partially or completely from a number of organisms including the yeast *Saccharomyces cerevisiae* (Labbe-Bois and Labbe, In *Biosynthesis of Heme and Chlorophyll*, E.H. Dailey, ed. McGraw Hill: New York, pp. 235-285 (1990)), barley etioplasts (Jacobs and Jacobs, *Biochem. J.* 244: 219 (1987)), and mouse liver (Dailey and Karr, *Biochem.* 26: 2697 (1987)). Genes encoding protox have been isolated from two prokaryotic organisms, *Escherichia coli* (Sasarman *et al.*, *Can. J. Microbiol.* 39: 1155 (1993)) and *Bacillus subtilis* (Dailey *et al.*, *J. Biol. Chem.* 269: 813 (1994)). These genes share no sequence similarity; neither do their predicted protein products share any amino acid sequence identity. The *E. coli* protein is approximately 21 kDa, and associates with the cell membrane. The *B. subtilis* protein is 51 kDa, and is a soluble, cytoplasmic activity.

Protox encoding genes have now also been isolated from humans (*see Nishimura et al.*, *J. Biol. Chem.* 270(14): 8076-8080 (1995) and plants (International application no. PCT/IB95/00452 filed June 8, 1995, published Dec. 21, 1995 as WO 95/34659).

II. The Protox Gene as a Herbicide Target

The use of herbicides to control undesirable vegetation such as weeds or plants in crops has become an almost universal practice. The relevant market exceeds a billion dollars annually. Despite this extensive use, weed control remains a significant and costly problem for farmers.

Effective use of herbicides requires sound management. For instance, time and method of application and stage of weed plant development are critical to getting good weed control with herbicides. Since various weed species are resistant to herbicides, the production of effective herbicides becomes increasingly important. Novel herbicides can now be discovered using high-throughput screens that implement recombinant DNA technology. Metabolic enzymes essential to plant growth and development can be recombinantly produced through standard molecular biological techniques and utilized as herbicide targets in screens for novel inhibitors of the enzymes' activity. The novel inhibitors discovered through such screens may then be used as herbicides to control undesirable vegetation.

Unfortunately, herbicides that exhibit greater potency, broader weed spectrum and more rapid degradation in soil can also have greater crop phytotoxicity. One solution applied to this problem has been to develop crops that are resistant or tolerant to herbicides. Crop hybrids or varieties resistant to the herbicides allow for the use of the herbicides without attendant risk of damage to the crop. Development of resistance can allow application of a herbicide to a crop where its use was previously precluded or limited (e.g. to pre-emergence use) due to sensitivity of the crop to the herbicide. For example, U.S. Patent No. 4,761,373, incorporated herein by reference, is directed to plants resistant to various imidazolinone or sulfonamide herbicides. The resistance is conferred by an altered acetohydroxyacid synthase (AHAS) enzyme. U.S. Patent No. 4,975,374, incorporated herein by reference, relates to plant cells and plants containing a gene encoding a mutant glutamine synthetase (GS) resistant to inhibition by herbicides that were known to inhibit GS, e.g. phosphinothricin and methionine sulfoximine. U.S. Patent No. 5,013,659, incorporated herein by reference, is directed to plants that express a mutant acetolactate synthase (ALS) that renders the plants resistant to inhibition by sulfonylurea herbicides. U.S. Patent No. 5,162,602, incorporated herein by reference, discloses plants tolerant to inhibition by cyclohexanedione and aryloxyphenoxypropanoic acid herbicides. The tolerance is conferred by an altered acetyl coenzyme A carboxylase (ACCase). U.S. Patent No. 5,554,798, incorporated herein by reference, discloses transgenic glyphosate resistant maize plants, which tolerance is conferred by an altered 5-enolpyruvyl-3-phosphoshikimate (EPSP) synthase gene.

The protox enzyme serves as the target for a variety of herbicidal compounds. The herbicides that inhibit protox include many different structural classes of molecules (Duke *et al.*, *Weed Sci.* 39: 465 (1991); Nandihalli *et al.*, *Pesticide Biochem. Physiol.* 43: 193 (1992); Matringe *et al.*, *FEBS Lett.* 245: 35 (1989); Yanase and Andoh, *Pesticide Biochem. Physiol.* 35: 70 (1989)). These herbicidal compounds include the diphenylethers {e.g. acifluorfen, 5-[2-chloro-4-(trifluoromethyl)phenoxy]-2-nitrobenzoic acid; its methyl ester; or oxyfluorfen, 2-chloro-1-(3-ethoxy-4-nitrophenoxy)-4-(trifluorobenzene)}, oxidiazoles, (e.g. oxidiazon, 3-[2,4-dichloro-5-(1-methylethoxy)phenyl]-5-(1,1-dimethylethyl)-1,3,4-oxadiazol-2-(3*H*)-one), cyclic imides (e.g. S-23142, *N*-(4-chloro-2-fluoro-5-propargyloxyphenyl)-3,4,5,6-tetrahydrophthalimide; chlorophthalim, *N*-(4-chlorophenyl)-3,4,5,6-tetrahydrophthalimide), phenyl pyrazoles (e.g. TNPP-ethyl, ethyl 2-[1-(2,3,4-trichlorophenyl)-4-nitropyrazolyl-5-oxy]propionate; M&B 39279), pyridine derivatives (e.g. LS 82-556), and phenopylate and its *O*-phenylpyrrolidino- and piperidinocarbamate analogs. Many of these compounds competitively inhibit the normal reaction catalyzed by the enzyme, apparently acting as substrate analogs.

Typically, the inhibitory effect on protox is determined by measuring fluorescence at about 622 to 635 nm, after excitation at about 395 to 410 nm (see, e.g. Jacobs and Jacobs, *Enzyme* 28: 206 (1982); Sherman *et al.*, *Plant Physiol.* 97: 280 (1991)). This assay is based on the fact that protoporphyrin IX is a fluorescent pigment, and protoporphyrinogen IX is nonfluorescent.

The predicted mode of action of protox-inhibiting herbicides involves the accumulation of protoporphyrinogen IX in the chloroplast. This accumulation is thought to lead to leakage of protoporphyrinogen IX into the cytosol where it is oxidized by a peroxidase activity to protoporphyrin IX. When exposed to light, protoporphyrin IX can cause formation of singlet oxygen in the cytosol. This singlet oxygen can in turn lead to the formation of other reactive oxygen species, which can cause lipid peroxidation and membrane disruption leading to rapid cell death (Lee *et al.*, *Plant Physiol.* 102: 881 (1993)).

Not all protox enzymes are sensitive to herbicides that inhibit plant protox enzymes. Both of the protox enzymes encoded by genes isolated from *Escherichia coli* (Sasarman *et al.*, *Can. J. Microbiol.* 39: 1155 (1993)) and *Bacillus subtilis* (Dailey *et al.*, *J. Biol. Chem.* 269: 813 (1994)) are resistant to these herbicidal inhibitors. In addition, mutants of the unicellular alga *Chlamydomonas reinhardtii* resistant to the phenylimide herbicide S-23142 have been reported (Kataoka *et al.*, *J. Pesticide Sci.* 15: 449 (1990); Shibata *et al.*, In *Research in Photosynthesis*, Vol. III, N. Murata, ed. Kluwer: Netherlands. pp. 567-570 (1992)). At least one of these mutants appears to have an altered protox activity that is

resistant not only to the herbicidal inhibitor on which the mutant was selected, but also to other classes of protox inhibitors (Oshio *et al.*, *Z. Naturforsch.* 48c: 339 (1993); Sato *et al.*, In *ACS Symposium on Porphyrin Pesticides*, S. Duke, ed. ACS Press: Washington, D.C. (1994)). A mutant tobacco cell line has also been reported that is resistant to the inhibitor S-21432 (Che *et al.*, *Z. Naturforsch.* 48c: 350 (1993)).

III. Plastid Transformation and Expression

Plastid transformation, in which genes are inserted by homologous recombination into some or all of the several thousand copies of the circular plastid genome present in each plant cell, takes advantage of the enormous copy number advantage over nuclear-expressed genes to permit expression levels that may exceed 10% of the total soluble plant protein. In addition, plastid transformation is desirable because in most plants plastid-encoded traits are not pollen transmissible; hence, potential risks of inadvertent transgene escape to wild relatives of transgenic plants is obviated. Plastid transformation technology is extensively described in U.S. Patent Nos. 5,451,513, 5,545,817, 5,545,818, and 5,576,198; in PCT Application Nos. WO 95/16783 and WO 97/32977; and in McBride *et al.*, *Proc. Natl. Acad. Sci. USA* 91: 7301-7305 (1994), all of which are incorporated herein by reference. Plastid transformation via biolistics was achieved initially in the unicellular green alga *Chlamydomonas reinhardtii* (Boynton *et al.* (1988) *Science* 240: 1534-1537, incorporated herein by reference) and this approach, using selection for *cis*-acting antibiotic resistance loci (spectinomycin/streptomycin resistance) or complementation of non-photosynthetic mutant phenotypes, was soon extended to *Nicotiana tabacum* (Svab *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87: 8526-8530, incorporated herein by reference).

The basic technique for tobacco chloroplast transformation involves the particle bombardment of leaf tissue or PEG-mediated uptake of plasmid DNA in protoplasts with regions of cloned plastid DNA flanking a selectable antibiotic resistance marker. The 1 to 1.5 kb flanking regions, termed "targeting sequences," facilitate homologous recombination with the plastid genome and thus allow the replacement or modification of specific regions of the 156 kb tobacco plastid genome. Initially, point mutations in the chloroplast 16S rDNA and *rps12* genes conferring resistance to spectinomycin and/or streptomycin were utilized as selectable markers for transformation (Svab, Z., Hajdukiewicz, P., and Maliga, P. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8526-8530; Staub, J. M., and Maliga, P. (1992) *Plant Cell* 4, 39-45, incorporated herein by reference). This resulted in stable homoplasmic transformants at a frequency of approximately one per 100 bombardments of target leaves. The presence of cloning sites between these markers allowed creation of a plastid targeting

vector for introduction of foreign genes (Staub, J.M., and Maliga, P., *EMBO J.* 12: 601-606 (1993), incorporated herein by reference). Substantial increases in transformation frequency were obtained by replacement of the recessive rRNA or r-protein antibiotic resistance genes with a dominant selectable marker, the bacterial *aadA* gene encoding the spectinomycin-detoxifying enzyme aminoglycoside-3'-adenyltransferase (Svab, Z., and Maliga, P. (1993) *Proc. Natl. Acad. Sci. USA* 90, 913-917, incorporated herein by reference). Previously, this marker had been used successfully for high-frequency transformation of the plastid genome of the green alga *Chlamydomonas reinhardtii* (Goldschmidt-Clermont, M. (1991) *Nucl. Acids Res.* 19, 4083-4089, incorporated herein by reference). Recently, plastid transformation of protoplasts from tobacco and the moss *Physcomitrella patens* has been attained using polyethylene glycol (PEG) mediated DNA uptake (O'Neill *et al.* (1993) *Plant J.* 3: 729-738; Koop *et al.* (1996) *Planta* 199: 193-201, both of which are incorporated herein by reference).

The present invention provides DNA molecules isolated from maize and comprising the coding sequence of an enzyme having protoporphyrinogen oxidase (protox) activity. In particular, the coding sequence of a mature maize protoporphyrinogen oxidase is set forth in SEQ ID NO: 44, and the corresponding amino acid sequence for a mature maize protoporphyrinogen oxidase enzyme is set forth in SEQ ID NO:45. The present invention further provides chimeric genes comprising such DNA molecules.

The present invention also provides DNA molecules isolated from wheat, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, and sugar cane encoding enzymes having protoporphyrinogen oxidase (protox) activity and chimeric genes comprising such DNA. Sequences of such DNA molecules are set forth in SEQ ID NOs: 9 (wheat), 11 (soybean), 15 (cotton), 17 (sugar beet), 19 (oilseed rape), 21 (rice), 23 (sorghum), and 36 (sugar cane).

The present invention also provides modified forms of plant protoporphyrinogen oxidase (protox) enzymes that are resistant to compounds that inhibit unmodified naturally occurring plant protox enzymes, and DNA molecules coding for such inhibitor-resistant plant protox enzymes. Thus, in one aspect the present invention provides a DNA molecule encoding a plant protox enzyme that is capable of being incorporated into a DNA construct used to transform a plant containing wild-type, herbicide-sensitive protox, wherein the DNA molecule has at least one point mutation relative to a wild-type DNA molecule encoding plant protox such that upon transformation with the DNA construct the plant contains the DNA molecule, which renders the plant resistant to the application of a herbicide that inhibits naturally occurring plant protox. The present invention includes chimeric genes and

modified forms of naturally occurring protox genes that can express the inhibitor-resistant plant protox enzymes in plants.

Genes encoding inhibitor-resistant plant protox enzymes can be used to confer resistance to protox-inhibitory herbicides in whole plants and as a selectable marker in plant cell transformation methods. Accordingly, the present invention also includes plants, including the descendants thereof, plant tissues and plant seeds containing plant expressible genes encoding these modified protox enzymes. These plants, plant tissues and plant seeds are resistant to protox-inhibitors at levels that normally are inhibitory to the naturally occurring protox activity in the plant. Plants encompassed by the invention especially include those that would be potential targets for protox inhibiting herbicides, particularly agronomically important crops such as maize and other cereal crops such as barley, wheat, sorghum, rye, oats, turf and forage grasses, millet and rice. Also comprised are other crop plants such as sugar cane, soybean, cotton, sugar beet, oilseed rape and tobacco.

The present invention accordingly provides a method for selecting plant cells transformed with a DNA molecule of the invention that encodes a herbicide-tolerant form of plant protox. The method comprises introducing the DNA molecule into plant cells whose growth is sensitive to inhibition by herbicides to which the protox encoded by the DNA molecule is resistant, thus forming a transformed plant cell. The transformed plant cell whose growth is resistant to the selected herbicide is identified by selection at a herbicide concentration that inhibits the growth of untransformed plant cells.

The present invention is directed further to methods for the production of plants, including plant material, such as for example plant tissues, protoplasts, cells, calli, organs, plant seeds, embryos, pollen, egg cells, zygotes, together with any other propagating material and plant parts, such as for example flowers, stems, fruits, leaves, roots originating in transgenic plants or their progeny previously transformed by means of the process of the invention, which produce an inhibitor-resistant form of the plant protox enzyme provided herein. Such plants may be stably transformed with a structural gene encoding the resistant protox, or prepared by direct selection techniques whereby herbicide resistant lines are isolated, characterized and developed.

In another aspect, the present invention is directed to a method for controlling unwanted vegetation growing at a locus where a herbicide-tolerant, agronomically useful plant, which is transformed with a DNA molecule according to the present invention that encodes a herbicide-tolerant form of plant protox, has been cultivated. The method

comprises applying to the locus to be protected an effective amount of herbicide that inhibits naturally occurring protox activity.

The present invention is further directed to probes and methods for detecting the presence of genes encoding inhibitor-resistant forms of the plant protox enzyme and quantitating levels of inhibitor-resistant protox transcripts in plant tissue. These methods may be used to identify or screen for plants or plant tissue containing and/or expressing a gene encoding an inhibitor-resistant form of the plant protox enzyme.

The present invention also relates to plastid transformation and to the expression of DNA molecules in a plant plastid. In a preferred embodiment, a native plant protox enzyme or a modified plant protox enzyme is expressed in plant plastids to obtain herbicide resistant plants.

In a further embodiment, the present invention is directed to a chimeric gene comprising: (a) a DNA molecule isolated from a plant, which in its native state encodes a polypeptide that comprises a plastid transit peptide, and a mature enzyme that is natively targeted to a plastid of the plant by the plastid transit peptide, wherein the DNA molecule is modified such that it does not encode a functional plastid transit peptide; and (b) a promoter capable of expressing the DNA molecule in a plastid, wherein the promoter is operatively linked to the DNA molecule. The DNA molecule may be modified in that at least a portion of the native plastid transit peptide coding sequence is absent from the DNA molecule. Alternatively, the DNA molecule may be modified in that one or more nucleotides of the native plastid transit peptide coding sequence are mutated, thereby rendering an encoded plastid transit peptide nonfunctional. The present invention also relates to plants homoplasmic for chloroplast genomes containing such chimeric genes. In a preferred embodiment, the DNA molecule encodes an enzyme that is naturally inhibited by a herbicidal compound. In this case, such plants are resistant to a herbicide that naturally inhibits the enzyme encoded by a DNA molecule according to the present invention.

The present invention is also directed to plants made resistant to a herbicide by transforming their plastid genome with a DNA molecule according to the present invention and to methods for obtaining such plants. In a preferred embodiment, the DNA molecule encodes an enzyme that is naturally inhibited by a herbicidal compound. In a more preferred embodiment, the DNA molecule encodes an enzyme having protoporphyrinogen oxidase (protox) activity, which is modified so that it that confers resistance to protox inhibitors. A further embodiment of the present invention is directed to a method for controlling the growth of undesired vegetation, which comprises applying to a population of the above-described plants an effective amount of an inhibitor of the enzyme.

The present invention also provides a novel method for selecting a transplastomic plant cell, comprising the steps of: introducing the above-described chimeric gene into the plastome of a plant cell; expressing the encoded enzyme in the plastids of the plant cell; and selecting a cell that is resistant to a herbicidal compound that naturally inhibits the activity of the enzyme, whereby the resistant cell comprises transformed plastids. In a preferred embodiment, the enzyme is naturally inhibited by a herbicidal compound and the transgenic plant is able to grow on an amount of the herbicidal compound that naturally inhibits the activity of the enzyme. In a further preferred embodiment, the enzyme has protoporphyrinogen oxidase (protox) activity and is modified so that it confers resistance to protox inhibitors.

Brief description of the Sequence Listing:

- SEQ ID NO:1: DNA coding sequence for an *Arabidopsis thaliana* protox-1 protein.
- SEQ ID NO:2: *Arabidopsis* protox-1 amino acid sequence encoded by SEQ ID NO:1.
- SEQ ID NO:3: DNA coding sequence for an *Arabidopsis thaliana* protox-2 protein.
- SEQ ID NO:4: *Arabidopsis* protox-2 amino acid sequence encoded by SEQ ID NO:3.
- SEQ ID NO:5: Partial DNA coding sequence for a maize protox-1 protein.
- SEQ ID NO:6: Partial maize protox-1 amino acid sequence encoded by SEQ ID NO:5.
- SEQ ID NO:7: DNA coding sequence for a maize protox-2 protein.
- SEQ ID NO:8: Maize protox-2 amino acid sequence encoded by SEQ ID NO:7.
- SEQ ID NO:9: Partial DNA coding sequence for a wheat protox-1 protein.
- SEQ ID NO:10: Partial wheat protox-1 amino acid sequence encoded by SEQ ID NO:9.
- SEQ ID NO:11: DNA coding sequence for a soybean protox-1 protein.
- SEQ ID NO:12: Soybean protox-1 protein encoded by SEQ ID NO:11.
- SEQ ID NO:13: Promoter sequence from *Arabidopsis thaliana* protox-1 gene.
- SEQ ID NO:14: Promoter sequence from maize protox-1 gene.
- SEQ ID NO:15: DNA coding sequence for a cotton protox-1 protein.
- SEQ ID NO:16: Cotton protox-1 amino acid sequence encoded by SEQ ID NO:15.
- SEQ ID NO:17: DNA coding sequence for a sugar beet protox-1 protein.
- SEQ ID NO:18: Sugar beet protox-1 amino acid sequence encoded by SEQ ID NO:17.
- SEQ ID NO:19: DNA coding sequence for an oilseed rape protox-1 protein.
- SEQ ID NO:20: Oilseed rape protox-1 amino acid sequence encoded by SEQ ID NO:19.
- SEQ ID NO:21: Partial DNA coding sequence for a rice protox-1 protein.
- SEQ ID NO:22: Partial rice protox-1 amino acid sequence encoded by SEQ ID NO:21.
- SEQ ID NO:23: Partial DNA coding sequence for a sorghum protox-1 protein.

SEQ ID NO:24: Partial sorghum protox-1 amino acid sequence encoded by SEQ ID NO:23.

SEQ ID NO:25: Maize protox-1 intron sequence.

SEQ ID NO:26: Promoter sequence from sugar beet protox-1 gene.

SEQ ID NO:27: Pclp_P1a - plastid *clpP* gene promoter top strand PCR primer.

SEQ ID NO:28: Pclp_P1b - plastid *clpP* gene promoter bottom strand PCR primer.

SEQ ID NO:29: Pclp_P2b - plastid *clpP* gene promoter bottom strand PCR primer.

SEQ ID NO:30: Trps16_P1a - plastid *rps16* gene top strand PCR primer.

SEQ ID NO:31: Trps16_p1b - plastid *rps16* gene bottom strand PCR primer.

SEQ ID NO:32: minpsb_U - plastid *psbA* gene top strand primer.

SEQ ID NO:33: minpsb_L - plastid *psbA* gene bottom strand primer.

SEQ ID NO:34: APRTXP1a - top strand PCR primer.

SEQ ID NO:35: APRTXP1b - bottom strand PCR primer.

SEQ ID NO:36: Partial DNA coding sequence for a sugar cane protox-1 protein.

SEQ ID NO:37: Partial sugar cane protox-1 amino acid sequence encoded by SEQ ID NO:36.

SEQ ID NO:38: Sub-sequence #1 (AP Δ_1 F).

SEQ ID NO:39: Sub-sequence #8 (YIGG Δ_8).

SEQ ID NO:40: Sub-sequence #12 (IGG Δ_{12}).

SEQ ID NO:41: Sub-sequence #13 (SWXL Δ_{13}).

SEQ ID NO:42: Sub-sequence #15 (G Δ_{15} XGL).

SEQ ID NO:43: Sub-sequence #17 (YV Δ_{17} G).

SEQ ID NO:44: DNA coding sequence for a mature maize protox-1 protein.

SEQ ID NO:45: Mature maize protox-1 amino acid sequence encoded by SEQ ID NO:44.

SEQ ID NO:46: Predicted coding sequence for a maize protox-1 polypeptide precursor.

SEQ ID NO:47: Predicted amino acid sequence of a maize protox-1 polypeptide precursor encoded by SEQ ID NO:46.

The following vector molecules have been deposited with Agricultural Research Service, Patent Culture Collection (NRRL), Northern Regional Research Center, 1815 North University Street, Peoria, Illinois 61604, U.S.A on the dates indicated below:

Wheat protox-1a, in the pBluescript SK vector, was deposited March 19, 1996, as pWDC-13 (NRRL #B21545).

Soybean protox-1, in the pBluescript SK vector, was deposited December 15, 1995 as pWDC-12 (NRRL #B-21516).

Cotton protox-1, in the pBluescript SK vector, was deposited July 1, 1996 as pWDC-15 (NRRL #B-21594).

Sugar beet protox-1, in the pBluescript SK vector, was deposited July 29, 1996, as pWDC-16 (NRRL #B-21595N).

Oilseed rape protox-1, in the pBluescript SK vector, was deposited August 23, 1996, as pWDC-17 (NRRL #B-21615).

Rice protox-1, in the pBluescript SK vector, was deposited December 6, 1996, as pWDC-18 (NRRL #B-21648).

Sorghum protox-1, in the pBluescript SK vector, was deposited December 6, 1996, as pWDC-19 (NRRL #B-21649).

Resistant mutant pAraC-2Cys, in the pMut-1 plasmid, was deposited on November 14, 1994 under the designation pWDC-7 with the Agricultural Research Culture Collection and given the deposit designation NRRL #21339N.

AraPT1Pro containing the *Arabidopsis* protox-1 promoter was deposited December 15, 1995, as pWDC-11 (NRRL #B-21515).

A plasmid containing the maize protox-1 promoter fused to the remainder of the maize protox-1 coding sequence was deposited March 19, 1996 as pWDC-14 (NRRL #B-21546).

A plasmid containing the sugar beet protox-1 promoter was deposited December 6, 1996, as pWDC-20 (NRRL #B-21650).

For clarity, certain terms used in the specification are defined and presented as follows:

Associated With / Operatively Linked: refers to two DNA sequences that are related physically or functionally. For example, a promoter or regulatory DNA sequence is said to be "associated with" a DNA sequence that codes for an RNA or a protein if the two sequences are operatively linked, or situated such that the regulator DNA sequence will affect the expression level of the coding or structural DNA sequence.

Chimeric Gene: a recombinant DNA sequence in which a promoter or regulatory DNA sequence is operatively linked to, or associated with, a DNA sequence that codes for an mRNA or which is expressed as a protein, such that the regulator DNA sequence is able to regulate transcription or expression of the associated DNA sequence. The regulator DNA sequence of the chimeric gene is not normally operatively linked to the associated DNA sequence as found in nature.

Coding DNA Sequence: a DNA sequence that is translated in an organism to produce a protein.

Corresponding To: in the context of the present invention, "corresponding to" means that when the amino acid sequences of various protox enzymes are aligned with each other, such as in Table 1A, the amino acids that "correspond to" certain enumerated positions in Table 1A are those that align with these positions in Table 1A, but that are not necessarily in these exact numerical positions relative to the particular protox enzyme's amino acid sequence. Likewise, when the amino acid sequence of a particular protox enzyme (for example, the soybean protox enzyme) is aligned with the amino acid sequence of a reference protox enzyme (for example, the *Arabidopsis* protox-1 sequence given in SEQ ID NO:2), the amino acids in the soybean protox sequence that "correspond to" certain enumerated positions of SEQ ID NO:2 are those that align with these positions of SEQ ID NO:2, but are not necessarily in these exact numerical positions of the soybean protox enzyme's amino acid sequence.

Herbicide: a chemical substance used to kill or suppress the growth of plants, plant cells, plant seeds, or plant tissues.

Heterologous DNA Sequence: a DNA sequence not naturally associated with a host cell into which it is introduced, including non-naturally occurring multiple copies of a naturally occurring DNA sequence.

Homologous DNA Sequence: a DNA sequence naturally associated with a host cell into which it is introduced.

Homoplasmic: refers to a plant, plant tissue or plant cell, wherein all of the plastids are genetically identical. In different tissues or stages of development, the plastids may take different forms, e.g., chloroplasts, proplastids, etioplasts, amyloplasts, chromoplasts, and so forth.

Inhibitor: a chemical substance that inactivates the enzymatic activity of a protein such as a biosynthetic enzyme, receptor, signal transduction protein, structural gene product, or transport protein that is essential to the growth or survival of the plant. In the context of the instant invention, an inhibitor is a chemical substance that inactivates the enzymatic activity of protox. The term "herbicide" is used herein to define an inhibitor when applied to plants, plant cells, plant seeds, or plant tissues.

Isolated: in the context of the present invention, an isolated nucleic acid molecule or an isolated enzyme is a nucleic acid molecule or enzyme that, by the hand of man, exists apart from its native environment and is therefore not a product of nature. An isolated

nucleic acid molecule or enzyme may exist in a purified form or may exist in a non-native environment such as, for example, a transgenic host cell.

Minimal Promoter: promoter elements, particularly a TATA element, that are inactive or that have greatly reduced promoter activity in the absence of upstream activation. In the presence of a suitable transcription factor, the minimal promoter functions to permit transcription.

Modified Enzyme Activity: enzyme activity different from that which naturally occurs in a plant (i.e. enzyme activity that occurs naturally in the absence of direct or indirect manipulation of such activity by man), which is tolerant to inhibitors that inhibit the naturally occurring enzyme activity.

Nucleic Acid Molecule: a linear segment of single- or double-stranded DNA or RNA that can be isolated from any source. In the context of the present invention, the nucleic acid molecule is preferably a segment of DNA.

Plant: refers to any plant or part of a plant at any stage of development. Therein are also included cuttings, cell or tissue cultures and seeds. As used in conjunction with the present invention, the term "plant tissue" includes, but is not limited to, whole plants, plant cells, plant organs, plant seeds, protoplasts, callus, cell cultures, and any groups of plant cells organized into structural and/or functional units.

Plastome: the genome of a plastid.

Protox-1: chloroplast protox.

Protox-2: mitochondrial protox.

Significant Increase: an increase in enzymatic activity that is larger than the margin of error inherent in the measurement technique, preferably an increase by about 2-fold or greater of the activity of the wild-type enzyme in the presence of the inhibitor, more preferably an increase by about 5-fold or greater, and most preferably an increase by about 10-fold or greater.

Substantially Similar: with respect to nucleic acids, a nucleic acid molecule that has at least 60 percent sequence identity with a reference nucleic acid molecule. In a preferred embodiment, a substantially similar DNA sequence is at least 80% identical to a reference DNA sequence; in a more preferred embodiment, a substantially similar DNA sequence is at least 90% identical to a reference DNA sequence; and in a most preferred embodiment, a substantially similar DNA sequence is at least 95% identical to a reference DNA sequence. A substantially similar nucleotide sequence typically hybridizes to a reference nucleic acid molecule, or fragments thereof, under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1%

SDS, at 50°C. With respect to proteins or peptides, a substantially similar amino acid sequence is an amino acid sequence that is at least 90% identical to the amino acid sequence of a reference protein or peptide and has substantially the same activity as the reference protein or peptide.

Tolerance / Resistance: the ability to continue normal growth or function when exposed to an inhibitor or herbicide.

Transformation: a process for introducing heterologous DNA into a cell, tissue, or plant. Transformed cells, tissues, or plants are understood to encompass not only the end product of a transformation process, but also transgenic progeny thereof.

Transit Peptide: a signal polypeptide that is translated in conjunction with a protein encoded by a DNA molecule, forming a polypeptide precursor. In the process of transport to a selected site within the cell, a chloroplast for example, the transit peptide can be cleaved from the remainder of the polypeptide precursor to provide an active or mature protein.

Transformed: refers to an organism such as a plant into which a heterologous DNA molecule has been introduced. The DNA molecule can be stably integrated into the genome of the plant, wherein the genome of the plant encompasses the nuclear genome, the plastid genome and the mitochondrial genome. In a transformed plant, the DNA molecule can also be present as an extrachromosomal molecule. Such an extrachromosomal molecule can be auto-replicating. A "non-transformed" plant refers to a wild-type organism, i.e., a plant, which does not contain the heterologous DNA molecule.

Transplastome: a transformed plastid genome.

Nucleotides are indicated by their bases by the following standard abbreviations: adenine (A), cytosine (C), thymine (T), and guanine (G). Amino acids are likewise indicated by the following standard abbreviations: alanine (ala; A), arginine (Arg; R), asparagine (Asn; N), aspartic acid (Asp; D), cysteine (Cys; C), glutamine (Gln; Q), glutamic acid (Glu; E), glycine (Gly; G), histidine (His; H), isoleucine (Ile; I), leucine (Leu; L), lysine (lys; K), methionine (Met; M), phenylalanine (Phe; F), proline (Pro; P), serine (Ser; S), threonine (Thr; T), tryptophan (Trp; W), tyrosine (Tyr; Y), and valine (Val; V). Furthermore (Xaa; X) represents any amino acid.

I. Plant Protox Coding Sequences

In one aspect, the inventors of the present invention have determined for the first time the NH₂ terminus of the mature maize protoporphyrin oxidase. To achieve this result, maize etioplasts were isolated and the protoporphyrin oxidase was immunoprecipitated from the etioplast extract using anti-protoporphyrin oxidase antibodies covalently bound to

Protein A Sepharose resin. The NH₂ terminus of the mature maize protoporphyrin oxidase was determined by microsequencing and found to be located at the alanine at position 36 in the amino acid sequence set forth in SEQ ID NO:47. This is an unexpected and surprising result since using the software program ChloroP (Nielsen, H., Engelbrecht, J., Brunak, S., & von Heijne, G. (1997) *Prot. Engineer.* 10, 1-6), the NH₂ terminus of the mature protein was predicted at the amino acid corresponding to the cysteine at position 35 in the amino acid sequence set forth in SEQ ID NO:47. The present invention therefore discloses a novel and non-obvious DNA molecule comprising the mature coding sequence of the maize protoporphyrin oxidase (protox). The nucleotide sequence of the coding sequence of the mature maize protoporphyrin oxidase (protox) is provided in SEQ ID NO:44 and the corresponding amino acid sequence is set forth in SEQ ID NO:45.

In another aspect, the present invention is directed to an isolated DNA molecule that encodes protoporphyrinogen oxidase (referred to herein as "protox"), the enzyme that catalyzes the oxidation of protoporphyrinogen IX to protoporphyrin IX, from wheat, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, and sugar cane. The partial DNA coding sequence and corresponding amino acid sequence for a wheat protox enzyme are provided as SEQ ID NOs:9 and 10, respectively. The DNA coding sequence and corresponding amino acid sequence for a soybean protox enzyme are provided as SEQ ID NOs:11 and 12, respectively. The DNA coding sequence and corresponding amino acid sequence for a cotton protox enzyme are provided as SEQ ID NOs:15 and 16, respectively. The DNA coding sequence and corresponding amino acid sequence for a sugar beet protox enzyme are provided as SEQ ID NOs:17 and 18, respectively. The DNA coding sequence and corresponding amino acid sequence for an oilseed rape protox enzyme are provided as SEQ ID NOs:19 and 20, respectively. The partial DNA coding sequence and corresponding amino acid sequence for a rice protox enzyme are provided as SEQ ID NOs:21 and 22, respectively. The partial DNA coding sequence and corresponding amino acid sequence for a sorghum protox enzyme are provided as SEQ ID NOs:23 and 24, respectively. The partial DNA coding sequence and corresponding amino acid sequence for a sugar cane protox enzyme are provided as SEQ ID NOs:36 and 37, respectively.

The DNA coding sequences and corresponding amino acid sequences for protox enzymes from *Arabidopsis thaliana* and maize are provided herein as SEQ ID NOs:1-4 (*Arabidopsis*) and SEQ ID NOs:5-8 (maize).

The invention therefore is directed to a DNA molecule encoding a protoporphyrinogen oxidase (protox) comprising a eukaryotic protox selected from the group consisting of a wheat protox enzyme, a soybean protox enzyme, a cotton protox

enzyme, a sugar beet protox enzyme, an oilseed rape protox enzyme, a rice protox enzyme, a sorghum protox enzyme, and a sugar cane protox enzyme.

Preferred within the scope of the invention are isolated DNA molecules encoding the protoporphyrinogen oxidase (protox) enzyme from dicotyledonous plants, but especially from soybean plants, cotton plants, sugar beet plants and oilseed rape plants, such as those given in SEQ ID NOS: 11, 15, 17 and 19. More preferred are isolated DNA molecules encoding the protoporphyrinogen oxidase (protox) enzyme from soybean, such as given in SEQ ID NO:11, and sugar beet, such as given in SEQ ID NO:17.

Also preferred are isolated DNA molecules encoding the protoporphyrinogen oxidase (protox) enzyme from monocotyledonous plants, but especially from wheat plants, rice plants, sorghum plants, and sugar cane plants, such as those given in SEQ ID NOS: 9, 21, 23, and 36. More preferred are isolated DNA molecules encoding the protoporphyrinogen oxidase (protox) enzyme from wheat such as given in SEQ ID NO:9.

In another aspect, the present invention is directed to isolated DNA molecules encoding the protoporphyrinogen oxidase (protox) enzyme protein from a dicotyledonous plant, wherein the protein comprises the amino acid sequence selected from the group consisting of SEQ ID NOs: 12, 16, 18 and 20. Further comprised are isolated DNA molecules encoding the protoporphyrinogen oxidase (protox) enzyme protein from a monocotyledonous plant, wherein the protein comprises the amino acid sequence selected from the group consisting of SEQ ID NOs: 10, 22, 24, and 37. More preferred is an isolated DNA molecule encoding the protoporphyrinogen oxidase (protox) enzyme wherein the protein comprises the amino acid sequence from wheat such as given in SEQ ID NO:10. More preferred is an isolated DNA molecule encoding the protoporphyrinogen oxidase (protox) enzyme wherein the protein comprises the amino acid sequence from soybean, such as given in SEQ ID NO:12 and sugar beet, such as given in SEQ ID NO:18.

Using the information provided by the present invention, the DNA coding sequence for the protoporphyrinogen oxidase (protox) enzyme from any eukaryotic organism may be obtained using standard methods.

In another aspect, the present invention is directed to an isolated DNA molecule that encodes a wheat protox enzyme and that comprises a nucleotide sequence that hybridizes to the coding sequence shown in SEQ ID NO:9 under the following hybridization and wash conditions:

- (a) hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50° C; and
- (b) wash in 2X SSC, 1% SDS at 50° C.

In yet another aspect, the present invention is directed to an isolated DNA molecule that encodes a soybean protox enzyme and that comprises a nucleotide sequence that hybridizes to the coding sequence shown in SEQ ID NO:11 under the following hybridization and wash conditions:

- (a) hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50° C; and
- (b) wash in 2X SSC, 1% SDS at 50° C.

In still another aspect, the present invention is directed to an isolated DNA molecule that encodes a cotton protox enzyme and that comprises a nucleotide sequence that hybridizes to the coding sequence shown in SEQ ID NO:15 under the following hybridization and wash conditions:

- (a) hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50° C; and
- (b) wash in 2X SSC, 1% SDS at 50° C.

In another aspect, the present invention is directed to an isolated DNA molecule that encodes a sugar beet protox enzyme and that comprises a nucleotide sequence that hybridizes to the coding sequence shown in SEQ ID NO:17 under the following hybridization and wash conditions:

- (a) hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50° C; and
- (b) wash in 2X SSC, 1% SDS at 50° C.

In another aspect, the present invention is directed to an isolated DNA molecule that encodes an oilseed rape protox enzyme and that comprises a nucleotide sequence that hybridizes to the coding sequence shown in SEQ ID NO:19 under the following hybridization and wash conditions:

- (a) hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50° C; and
- (b) wash in 2X SSC, 1% SDS at 50° C.

In another aspect, the present invention is directed to an isolated DNA molecule that encodes a rice protox enzyme and that comprises a nucleotide sequence that hybridizes to the coding sequence shown in SEQ ID NO:21 under the following hybridization and wash conditions:

- (a) hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50° C; and
- (b) wash in 2X SSC, 1% SDS at 50° C.

In another aspect, the present invention is directed to an isolated DNA molecule that encodes a sorghum protox enzyme and that comprises a nucleotide sequence that hybridizes to the coding sequence shown in SEQ ID NO:23 under the following hybridization and wash conditions:

- (a) hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50° C; and
- (b) wash in 2X SSC, 1% SDS at 50° C.

In another aspect, the present invention is directed to an isolated DNA molecule that encodes a sugar cane protox enzyme and that comprises a nucleotide sequence that hybridizes to the coding sequence shown in SEQ ID NO:36 under the following hybridization and wash conditions:

- (a) hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50° C; and
- (b) wash in 2X SSC, 1% SDS at 50° C.

The isolated eukaryotic protox sequences taught by the present invention may be manipulated according to standard genetic engineering techniques to suit any desired purpose. For example, the entire protox sequence or portions thereof may be used as probes capable of specifically hybridizing to protox coding sequences and messenger RNA's. To achieve specific hybridization under a variety of conditions, such probes include sequences that are unique among protox coding sequences and are preferably at least 10 nucleotides in length, and most preferably at least 20 nucleotides in length. Such probes may be used to amplify and analyze protox coding sequences from a chosen organism via the well known process of polymerase chain reaction (PCR). This technique may be useful to isolate additional protox coding sequences from a desired organism or as a diagnostic assay to determine the presence of protox coding sequences in an organism.

Factors that affect the stability of hybrids determine the stringency of the hybridization. One such factor is the melting temperature T_m , which can be easily calculated according to the formula provided in DNA PROBES, George H. Keller and Mark M. Manak, Macmillan Publishers Ltd, 1993, Section one: Molecular Hybridization Technology; page 8 ff.

The preferred hybridization temperature is in the range of about 25°C below the calculated melting temperature T_m and preferably in the range of about 12-15°C below the calculated melting temperature T_m and in the case of oligonucleotides in the range of about 5-10°C below the melting temperature T_m .

Comprised by the present invention are DNA molecules that hybridize to a DNA molecule according to the invention as defined hereinbefore, but preferably to an oligonucleotide probe obtainable from the DNA molecule comprising a contiguous portion of the sequence of the protoporphyrinogen oxidase (protox) enzyme at least 10 nucleotides in length, under moderately stringent conditions.

The invention further embodies the use of a nucleotide probe capable of specifically hybridizing to a plant protox gene or mRNA of at least 10 nucleotides length in a polymerase chain reaction (PCR).

In a further embodiment, the present invention provides probes capable of specifically hybridizing to a eukaryotic DNA sequence encoding a protoporphyrinogen oxidase activity or to the respective mRNA and methods for detecting the DNA sequences in eukaryotic organisms using the probes according to the invention.

Protox specific hybridization probes may also be used to map the location of the native eukaryotic protox gene(s) in the genome of a chosen organism using standard techniques based on the selective hybridization of the probe to genomic protox sequences. These techniques include, but are not limited to, identification of DNA polymorphisms identified or contained within the protox probe sequence, and use of such polymorphisms to follow segregation of the protox gene relative to other markers of known map position in a mapping population derived from self fertilization of a hybrid of two polymorphic parental lines (see e.g. Helentjaris *et al.*, *Plant Mol. Biol.* 5: 109 (1985). Sommer *et al.* *Biotechniques* 12: 82 (1992); D'Ovidio *et al.*, *Plant Mol. Biol.* 15: 169 (1990)). While any eukaryotic protox sequence is contemplated to be useful as a probe for mapping protox genes from any eukaryotic organism, preferred probes are those protox sequences from organisms more closely related to the chosen organism, and most preferred probes are those protox sequences from the chosen organism. Mapping of protox genes in this manner is contemplated to be particularly useful in plants for breeding purposes. For instance, by knowing the genetic map position of a mutant protox gene that confers herbicide resistance, flanking DNA markers can be identified from a reference genetic map (see, e.g., Helentjaris, *Trends Genet.* 3: 217 (1987)). During introgression of the herbicide resistance trait into a new breeding line, these markers can then be used to monitor the extent of protox-linked flanking chromosomal DNA still present in the recurrent parent after each round of back-crossing.

Protox specific hybridization probes may also be used to quantitate levels of protox mRNA in an organism using standard techniques such as Northern blot analysis. This technique may be useful as a diagnostic assay to detect altered levels of protox expression

that may be associated with particular adverse conditions such as autosomal dominant disorder in humans characterized by both neuropsychiatric symptoms and skin lesions, which are associated with decreased levels of protox activity (Brenner and Bloomer, *New Engl. J. Med.* 302: 765 (1980)).

A further embodiment of the invention is a method of producing a DNA molecule comprising a DNA portion encoding a protein having protoporphyrinogen oxidase (protox) enzyme activity comprising:

- (a) preparing a nucleotide probe capable of specifically hybridizing to a plant protox gene or mRNA, wherein the probe comprises a contiguous portion of the coding sequence for a protox protein from a plant of at least 10 nucleotides length;
- (b) probing for other protox coding sequences in populations of cloned genomic DNA fragments or cDNA fragments from a chosen organism using the nucleotide probe prepared according to step (a); and
- (c) isolating and multiplying a DNA molecule comprising a DNA portion encoding a protein having protoporphyrinogen oxidase (protox) enzyme activity.

A further embodiment of the invention is a method of isolating a DNA molecule from any plant comprising a DNA portion encoding a protein having protoporphyrinogen oxidase (protox) enzyme activity.

- (a) preparing a nucleotide probe capable of specifically hybridizing to a plant protox gene or mRNA, wherein the probe comprises a contiguous portion of the coding sequence for a protox protein from a plant of at least 10 nucleotides length;
- (b) probing for other protox coding sequences in populations of cloned genomic DNA fragments or cDNA fragments from a chosen organism using the nucleotide probe prepared according to step (a); and
- (c) isolating a DNA molecule comprising a DNA portion encoding a protein having protoporphyrinogen oxidase (protox) enzyme activity.

The invention further comprises a method of producing an essentially pure DNA sequence coding for a protein exhibiting protoporphyrinogen oxidase (protox) enzyme activity, which method comprises:

- (a) preparing a genomic or a cDNA library from a suitable source organism using an appropriate cloning vector;
- (b) hybridizing the library with a probe molecule; and

- (c) identifying positive hybridizations of the probe to the DNA clones from the library that is clones potentially containing the nucleotide sequence corresponding to the amino acid sequence for protoporphyrinogen oxidase (protox).

The invention further comprises a method of producing an essentially pure DNA sequence coding for a protein exhibiting protoporphyrinogen oxidase (protox) enzyme activity, which method comprises:

- (a) preparing total DNA from a genomic or a cDNA library;
- (b) using the DNA of step (a) as a template for PCR reaction with primers representing low degeneracy portions of the amino acid sequence of protoporphyrinogen oxidase (protox).

A further object of the invention is an assay to identify inhibitors of protoporphyrinogen oxidase (protox) enzyme activity that comprises:

- (a) incubating a first sample of protoporphyrinogen oxidase (protox) and its substrate;
- (b) measuring an uninhibited reactivity of the protoporphyrinogen oxidase (protox) from step (a);
- (c) incubating a first sample of protoporphyrinogen oxidase (protox) and its substrate in the presence of a second sample comprising an inhibitor compound;
- (d) measuring an inhibited reactivity of the protoporphyrinogen oxidase (protox) enzyme from step (c); and
- (e) comparing the inhibited reactivity to the uninhibited reactivity of protoporphyrinogen oxidase (protox) enzyme.

A further object of the invention is an assay to identify inhibitor-resistant protoporphyrinogen oxidase (protox) mutants that comprises:

- (a) incubating a first sample of protoporphyrinogen oxidase (protox) enzyme and its substrate in the presence of a second sample comprising a protoporphyrinogen oxidase (protox) enzyme inhibitor;
- (b) measuring an unmutated reactivity of the protoporphyrinogen oxidase (protox) enzyme from step (a);
- (c) incubating a first sample of a mutated protoporphyrinogen oxidase (protox) enzyme and its substrate in the presence of a second sample comprising protoporphyrinogen oxidase (protox) enzyme inhibitor;

- (d) measuring a mutated reactivity of the mutated protoporphyrinogen oxidase (protox) enzyme from step (c); and
- (e) comparing the mutated reactivity to the unmutated reactivity of the protoporphyrinogen oxidase (protox) enzyme.

A further object of the invention is a protox enzyme inhibitor obtained by a method according to the invention.

For recombinant production of the enzyme in a host organism, the protox coding sequence may be inserted into an expression cassette designed for the chosen host and introduced into the host where it is recombinantly produced. The choice of specific regulatory sequences such as promoter, signal sequence, 5' and 3' untranslated sequences, and enhancer, is within the level of skill of the routineer in the art. The resultant molecule, containing the individual elements linked in proper reading frame, may be inserted into a vector capable of being transformed into the host cell. Suitable expression vectors and methods for recombinant production of proteins are well known for host organisms such as *E. coli* (see, e.g. Studier and Moffatt, *J. Mol. Biol.* 189: 113 (1986); Brosius, *DNA* 8: 759 (1989)), yeast (see, e.g., Schneider and Guarente, *Meth. Enzymol.* 194: 373 (1991)) and insect cells (see, e.g., Luckow and Summers, *Bio/Technol.* 6: 47 (1988)). Specific examples include plasmids such as pBluescript (Stratagene, La Jolla, CA), pFLAG (International Biotechnologies, Inc., New Haven, CT), pTrcHis (Invitrogen, La Jolla, CA), and baculovirus expression vectors, e.g., those derived from the genome of *Autographica californica* nuclear polyhedrosis virus (AcMNPV). A preferred baculovirus/insect system is pVI11392/Sf21 cells (Invitrogen, La Jolla, CA).

In a preferred embodiment, a coding sequence encoding a mature protox enzyme is inserted into an appropriate expression cassette and the mature protox enzyme is recombinantly produced in a desired host cell. Alternatively, a coding sequence encoding a protox polypeptide precursor comprising an entire plastid transit peptide or a portion of the plastid transit peptide is inserted into an appropriate expression cassette, and the protox polypeptide precursor is recombinantly produced in the host cell. The transit peptide is then cleaved to release the mature protein. The cleavage is for example carried out by in-vitro digestion of the polypeptide precursor, for example using a plastid transit peptide peptidase.

Recombinantly produced eukaryotic protox enzyme is useful for a variety of purposes. For example, it may be used to supply protox enzymatic activity *in vitro*. It may also be used in an *in vitro* assay to screen known herbicidal chemicals whose target has not been identified to determine if they inhibit protox. Such an *in vitro* assay may also be used as a more general screen to identify chemicals that inhibit protox activity and that are

therefore herbicide candidates. Recombinantly produced eukaryotic protox enzyme may also be used in an assay to identify inhibitor-resistant protox mutants (see International application no. PCT/IB95/00452 filed June 8, 1995, published Dec. 21, 1995 as WO 95/34659, incorporated by reference herein in its entirety). Alternatively, recombinantly produced protox enzyme may be used to further characterize its association with known inhibitors in order to rationally design new inhibitory herbicides as well as herbicide tolerant forms of the enzyme.

II. Inhibitor Resistant Plant Protox Enzymes

In another aspect, the present invention teaches modifications that can be made to the amino acid sequence of any eukaryotic protoporphyrinogen oxidase (referred to herein as "protox") enzyme to yield an inhibitor-resistant form of this enzyme. Preferably, the eukaryotic protox enzyme is a plant protox enzyme. The present invention is directed to inhibitor-resistant protox enzymes having the modifications taught herein, to DNA molecules encoding these modified enzymes, and to chimeric genes capable of expressing these modified enzymes in plants.

The present invention is thus directed to an isolated DNA molecule encoding a modified eukaryotic protoporphyrinogen oxidase (protox) having at least one amino acid modification, wherein the amino acid modification has the property of conferring resistance to a protox inhibitor, that is wherein the modified protox is tolerant to an inhibitor in amounts that inhibit the naturally occurring eukaryotic protox. As used herein "inhibit" refers to a reduction in enzymatic activity observed in the presence of a subject compound compared to the level of activity observed in the absence of the subject compound, wherein the percent-level of reduction is preferably at least 10%, more preferably at least 50%, and most preferably at least 90%.

Preferred is a DNA molecule encoding a modified eukaryotic protoporphyrinogen oxidase (protox) that is a plant protox, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Even more preferred is a protox selected from the group consisting of an *Arabidopsis* protox enzyme, a maize protox enzyme, a wheat protox enzyme, a soybean protox enzyme, a cotton protox enzyme, a sugar beet protox enzyme, an oilseed rape protox enzyme, a rice protox enzyme, a sorghum protox enzyme, and a sugar cane protox enzyme having at least one amino acid modification, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity.

As used herein, the expression "substantially conserved amino acid sequences" refers to regions of amino acid homology between polypeptides comprising protox enzymes from different sources. In the present invention, seventeen substantially conserved amino acid sub-sequences, designated 1-19 respectively, are shown in Table 1B. One skilled in the art could align the amino acid sequences of protox enzymes from different sources, as has been done in Table 1A, to identify the sub-sequences therein that make up the substantially conserved amino acid sequences defined herein. Stated another way, a given sub-sequence from one source "corresponds to" a homologous subsequence from a different source. The skilled person could then determine whether the identified sub-sequences have the characteristics disclosed and claimed in the present application.

Therefore, a preferred embodiment of the present invention is directed to a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme having protoporphyrinogen oxidase (protox) activity, wherein the nucleic acid molecule is capable of being incorporated into a nucleic acid construct used to transform a plant containing wild-type, herbicide-sensitive protox, wherein the nucleotide sequence has at least one point mutation relative to a wild-type nucleotide sequence encoding plant protox, such that upon transformation with the nucleic acid construct the plant is rendered herbicide-tolerant.

More particularly, a preferred embodiment of the present invention is directed to a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises at least one of the following amino acid sub-sequences:

- (a) $AP\Delta_1F$, wherein Δ_1 is an amino acid other than arginine;
- (b) $F\Delta_2S$, wherein Δ_2 is an amino acid other than cysteine;
- (c) $Y\Delta_3G$, wherein Δ_3 is an amino acid other than alanine;
- (d) $A\Delta_4D$, wherein Δ_4 is an amino acid other than glycine;
- (e) $Y\Delta_5P$, wherein Δ_5 is an amino acid other than proline;
- (f) $P\Delta_6A$, wherein Δ_6 is an amino acid other than valine;
- (g) Δ_7IG , wherein Δ_7 is an amino acid other than tyrosine;
- (h) $YIGG\Delta_8$, wherein Δ_8 is an amino acid other than alanine or serine;
- (i) $A\Delta_9P$, wherein Δ_9 is an amino acid other than isoleucine; and
- (j) $G\Delta_{10}A$, wherein Δ_{10} is an amino acid other than valine;
- (k) $KA\Delta_{18}F$, wherein Δ_{18} is an amino acid other than alanine;

(I) $Q\Delta_{19}H$, wherein Δ_{19} is an amino acid other than leucine.

(Table 1B; sub-sequences 1-10 and 18-19).

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $AP\Delta_1F$, wherein Δ_1 is an amino acid other than arginine. Most preferably, Δ_1 is cysteine or leucine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $F\Delta_2S$, wherein Δ_2 is an amino acid other than cysteine. Most preferably, Δ_2 is phenylalanine, leucine, or lysine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $Y\Delta_3G$, wherein Δ_3 is an amino acid other than alanine. Most preferably, Δ_3 is valine, threonine, leucine, cysteine, or isoleucine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $A\Delta_4D$, wherein Δ_4 is an amino acid other than glycine. Most preferably, Δ_4 is serine or leucine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence Δ_5 is an amino acid other than proline. Most preferably, Δ_5 is serine or histidine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring

protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $P\Delta_6A$, wherein Δ_6 is an amino acid other than valine. Most preferably, Δ_6 is leucine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence Δ_7IG , wherein Δ_7 is an amino acid other than tyrosine. Most preferably, Δ_7 is cysteine, isoleucine, leucine, threonine, methionine, valine, alanine, histidine, or arginine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $F\Delta_8S$, wherein Δ_8 is an amino acid other than alanine or serine. Most preferably, Δ_8 is proline.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $A\Delta_9P$, wherein Δ_9 is an amino acid other than isoleucine. Most preferably, Δ_9 is threonine, histidine, glycine, or asparagine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $G\Delta_{10}A$, wherein Δ_{10} is an amino acid other than valine. Most preferably, Δ_{10} is alanine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $K\Delta_{18}F$, wherein Δ_{18} is an amino acid other than alanine. Most preferably, Δ_{18} is threonine or valine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring

protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $Q\Delta_{19}H$, wherein Δ_{19} is an amino acid other than leucine. Most preferably, Δ_{19} is serine.

Another preferred embodiment of the present invention is directed to a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises at least one of the following amino acid sub-sequences:

- (a) $AP\Delta_1F$, wherein Δ_1 is an amino acid other than arginine;
- (b) $F\Delta_2S$, wherein Δ_2 is an amino acid other than cysteine;
- (c) $Y\Delta_3G$, wherein Δ_3 is an amino acid other than alanine;
- (d) $A\Delta_4D$, wherein Δ_4 is an amino acid other than glycine;
- (e) $Y\Delta_5P$, wherein Δ_5 is an amino acid other than proline;
- (f) $P\Delta_6A$, wherein Δ_6 is an amino acid other than valine;
- (g) Δ_7IG , wherein Δ_7 is an amino acid other than tyrosine;
- (h) $YIGG\Delta_8$, wherein Δ_8 is an amino acid other than alanine or serine;
- (i) $A\Delta_9P$, wherein Δ_9 is an amino acid other than isoleucine; and
- (j) $G\Delta_{10}A$, wherein Δ_{10} is an amino acid other than valine
- (k) $KA\Delta_{18}F$, wherein Δ_{18} is an amino acid other than alanine;
- (l) $Q\Delta_{19}H$, wherein Δ_{19} is an amino acid other than leucine.

(Table 1B; sub-sequences 1-10 and 18-19), and wherein the modified enzyme further comprises at least one additional amino acid sub-sequence selected from the group consisting of:

- (m) $Q\Delta_{11}S$, wherein Δ_{11} is an amino acid other than proline;
- (n) $IGG\Delta_{12}$, wherein Δ_{12} is an amino acid other than threonine;
- (o) $SWXL\Delta_{13}$, wherein Δ_{13} is an amino acid other than serine;
- (p) $L\Delta_{14}Y$, wherein Δ_{14} is an amino acid other than asparagine; and
- (q) $G\Delta_{15}XGL$, wherein Δ_{15} is an amino acid other than tyrosine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $Y\Delta_3G$, wherein Δ_3 is an amino acid other than alanine, or the amino acid sub-sequence Δ_7IG ,

wherein Δ_7 is an amino acid other than tyrosine, and wherein the modified enzyme further comprises at least one additional amino acid sub-sequence selected from the group consisting of:

- (m) $Q\Delta_{11}S$, wherein Δ_{11} is an amino acid other than proline;
- (n) $IGG\Delta_{12}$, wherein Δ_{12} is an amino acid other than threonine;
- (o) $SWXL\Delta_{13}$, wherein Δ_{13} is an amino acid other than serine;
- (p) $L\Delta_{14}Y$, wherein Δ_{14} is an amino acid other than asparagine; and
- (q) $G\Delta_{15}XGL$, wherein Δ_{15} is an amino acid other than tyrosine.

Preferably, Δ_{11} is leucine, Δ_{12} is isoleucine or alanine, Δ_{13} is leucine, Δ_{14} is serine, and Δ_{15} is cysteine.

Another preferred embodiment of the present invention is directed to a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises: the amino acid sub-sequence Δ_7IG , wherein Δ_7 is an amino acid other than tyrosine; the amino acid sub-sequences $IGG\Delta_{12}$, wherein Δ_{12} is an amino acid other than threonine; and the amino acid sub-sequence $SWXL\Delta_{13}$, wherein Δ_{13} is an amino acid other than serine. Most preferably, Δ_7 is isoleucine, Δ_{12} is isoleucine, and Δ_{13} is leucine.

Yet another preferred embodiment of the present invention is directed to a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the nucleotide sequence is further characterized in that at least one of the following conditions is met:

- (a) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $AP\Delta_1F$, wherein Δ_1 is an amino acid other than arginine;
- (b) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $F\Delta_2S$, wherein Δ_2 is an amino acid other than cysteine;
- (c) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $Y\Delta_3G$, wherein Δ_3 is an amino acid other than alanine;
- (d) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $A\Delta_4D$, wherein Δ_4 is an amino acid other than glycine;

- (e) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $Y\Delta_5P$, wherein Δ_5 is an amino acid other than proline;
- (f) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $P\Delta_6A$, wherein Δ_6 is an amino acid other than valine;
- (g) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence Δ_7IG , wherein Δ_7 is an amino acid other than tyrosine;
- (h) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $YIGG\Delta_8$, wherein Δ_8 is an amino acid other than alanine or serine;
- (i) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $A\Delta_9P$, wherein Δ_9 is an amino acid other than isoleucine;
- (j) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $G\Delta_{10}A$, wherein Δ_{10} is an amino acid other than valine;
- (k) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $Y\Delta_3G$, wherein Δ_3 is an amino acid other than alanine, and the nucleic acid sequence also has a sequence that encodes one of the group consisting of:
 - (1) sub-sequence $Q\Delta_{11}S$, wherein Δ_{11} is an amino acid other than proline,
 - (2) sub-sequence $IGG\Delta_{12}$, wherein Δ_{12} is an amino acid other than threonine,
 - (3) sub-sequence $SWXL\Delta_{13}$, wherein Δ_{13} is an amino acid other than serine,
 - (4) sub-sequence $L\Delta_{14}Y$, wherein Δ_{14} is an amino acid other than asparagine,
 - and
 - (5) sub-sequence $G\Delta_{15}XGL$, wherein Δ_{15} is an amino acid other than tyrosine;
- (l) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence Δ_7IG , wherein Δ_7 is an amino acid other than tyrosine, and the nucleic acid sequence also has a sequence that encodes one of the group consisting of:
 - (1) sub-sequence $Q\Delta_{11}S$, wherein Δ_{11} is an amino acid other than proline,
 - (2) sub-sequence $IGG\Delta_{12}$, wherein Δ_{12} is an amino acid other than threonine,
 - (3) sub-sequence $SWXL\Delta_{13}$, wherein Δ_{13} is an amino acid other than serine,
 - (4) sub-sequence $L\Delta_{14}Y$, wherein Δ_{14} is an amino acid other than asparagine,
 - and
 - (5) sub-sequence $G\Delta_{15}XGL$, wherein Δ_{15} is an amino acid other than tyrosine;

- (m) the nucleic acid has a sequence that encodes amino acid sub-sequence $T\Delta_{16}G$, wherein Δ_{16} is an amino acid other than leucine, and the nucleic acid sequence also has a sequence that encodes amino acid sub-sequence $YV\Delta_{17}G$, wherein Δ_{17} is an amino acid other than alanine.
- (n) $KA\Delta_{18}F$, wherein Δ_{18} is an amino acid other than alanine;
- (o) $Q\Delta_{19}H$, wherein Δ_{19} is an amino acid other than leucine.

Preferably, said nucleic acid sequence has a sequence that encodes amino acid sub-sequence $T\Delta_{16}G$, wherein Δ_{16} is an amino acid other than leucine, and said nucleic acid sequence also has a sequence that encodes amino acid sub-sequence $YV\Delta_{17}G$, wherein Δ_{17} is an amino acid other than alanine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the arginine occurring at the position corresponding to amino acid 88 of SEQ ID NO:6 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is the DNA molecule wherein the arginine is replaced with a cysteine or a leucine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the cysteine occurring at the position corresponding to amino acid 159 of SEQ ID NO:6 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is the DNA molecule wherein the cysteine is replaced with a phenylalanine or lysine, most preferred, wherein the cysteine is replaced with a phenylalanine.

Also preferred is a DNA encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the isoleucine occurring at the position corresponding to amino acid 419 of SEQ ID NO:6 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule, wherein the isoleucine is replaced with a threonine, histidine, glycine or asparagine most preferred, wherein the isoleucine is replaced with a threonine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the alanine occurring at the position corresponding to amino acid 164 of SEQ ID NO:6 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally

occurring protox activity. Particularly preferred is a DNA molecule wherein the alanine is replaced with a threonine, leucine or valine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the glycine occurring at the position corresponding to amino acid 165 of SEQ ID NO:6 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the glycine is replaced with a serine or leucine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the tyrosine occurring at the position corresponding to amino acid 370 of SEQ ID NO:6 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the tyrosine is replaced with a isoleucine or methionine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the alanine occurring at the position corresponding to amino acid 175 of SEQ ID NO:6 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the alanine is replaced with a valine or threonine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the leucine occurring at the position corresponding to amino acid 337 of SEQ ID NO:6 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the leucine is replaced with a serine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the valine occurring at the position corresponding to amino acid 356 of SEQ ID NO:10 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the valine is replaced with a leucine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the serine occurring at the position corresponding to amino acid 421 of SEQ ID NO:10 is replaced with another amino acid,

wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the serine is replaced with a proline.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the valine occurring at the position corresponding to amino acid 502 of SEQ ID NO:10 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the valine is replaced with a alanine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the alanine occurring at the position corresponding to amino acid 211 of SEQ ID NO:10 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the alanine is replaced with a valine or threonine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the glycine occurring at the position corresponding to amino acid 212 of SEQ ID NO:10 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the glycine is replaced with a serine.

Also preferred is a DNA encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the isoleucine occurring at the position corresponding to amino acid 466 of SEQ ID NO:10 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the isoleucine is replaced with a threonine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the proline occurring at the position corresponding to amino acid 369 of SEQ ID NO:12 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the proline is replaced with a serine or histidine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the alanine occurring at the position corresponding to amino acid 226 of SEQ ID NO:12 is replaced with another amino acid,

wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule, wherein the alanine is replaced with a threonine or leucine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the valine occurring at the position corresponding to amino acid 517 of SEQ ID NO:12 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the valine is replaced with a alanine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the tyrosine occurring at the position corresponding to amino acid 432 of SEQ ID NO:12 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the tyrosine is replaced with a leucine or isoleucine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the proline occurring at the position corresponding to amino acid 365 of SEQ ID NO:16 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the proline is replaced with a serine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the tyrosine occurring at the position corresponding to amino acid 428 of SEQ ID NO:16 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the tyrosine is replaced with a cysteine, histidine or arginine.

Also preferred is a DNA encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the tyrosine occurring at the position corresponding to amino acid 449 of SEQ ID NO:18 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the tyrosine is replaced with a cysteine, leucine, isoleucine, valine or methionine.

The present invention is further directed to a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox having a first amino acid

substitution and a second amino acid substitution; the first amino acid substitution having the property of conferring resistance to a protox inhibitor; and the second amino acid substitution having the property of enhancing the resistance conferred by the first amino acid substitution. Preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox, wherein the plant is selected from the group consisting of maize, wheat, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, sugar cane, and *Arabidopsis*. More preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox, wherein the plant is selected from the group consisting of maize, wheat, soybean, sugar beet, and *Arabidopsis*.

Preferred is a DNA molecule wherein the second amino acid substitution occurs at a position selected from the group consisting of:

- (a) the position corresponding to the serine at amino acid 305 of SEQ ID NO:2;
- (b) the position corresponding to the threonine at amino acid 249 of SEQ ID NO:2;
- (c) the position corresponding to the proline at amino acid 118 of SEQ ID NO:2;
- (d) the position corresponding to the asparagine at amino acid 425 of SEQ ID NO:2;

and

- (e) the position corresponding to the tyrosine at amino acid 498 of SEQ ID NO:2.

Also preferred is a DNA molecule wherein the first amino acid substitution occurs at a position selected from the group consisting of:

- (a) the position corresponding to the arginine at amino acid 88 of SEQ ID NO:6;
- (b) the position corresponding to the alanine at amino acid 164 of SEQ ID NO:6;
- (c) the position corresponding to the glycine at amino acid 165 of SEQ ID NO:6;
- (d) the position corresponding to the tyrosine at amino acid 370 of SEQ ID NO:6;
- (e) the position corresponding to the cysteine at amino acid 159 of SEQ ID NO:6;
- (f) the position corresponding to the isoleucine at amino acid 419 of SEQ ID NO:6.
- (g) the position corresponding to the valine at amino acid 356 of SEQ ID NO:10;
- (h) the position corresponding to the serine at amino acid 421 of SEQ ID NO:10;
- (i) the position corresponding to the valine at amino acid 502 of SEQ ID NO:10;
- (j) the position corresponding to the alanine at amino acid 211 of SEQ ID NO:10;
- (k) the position corresponding to the glycine at amino acid 212 of SEQ ID NO:10;
- (l) the position corresponding to the isoleucine at amino acid 466 of SEQ ID NO:10;
- (m) the position corresponding to the proline at amino acid 369 of SEQ ID NO:12;
- (n) the position corresponding to the alanine at amino acid 226 of SEQ ID NO:12;
- (o) the position corresponding to the tyrosine at amino acid 432 of SEQ ID NO:12;
- (p) the position corresponding to the valine at amino acid 517 of SEQ ID NO:12;

- (q) the position corresponding to the tyrosine at amino acid 428 of SEQ ID NO:16;
 - (r) the position corresponding to the proline at amino acid 365 of SEQ ID NO:16;
 - (s) the position corresponding to the tyrosine at amino acid 449 of SEQ ID NO:18;
 - (t) the position corresponding to the alanine at amino acid 175 of SEQ ID NO:6;
- and

- (u) the position corresponding to the leucine at amino acid 337 of SEQ ID NO:6.

Particularly preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the plant protox comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 16, 18, 20, 22, and 37. Most preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox, wherein the plant protox comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 2, 4, 6, 8, 10, 12, and 18.

More preferred is a DNA molecule, wherein the first amino acid substitution occurs at a position selected from the group consisting of:

- (a) the position corresponding to the arginine at amino acid 88 of SEQ ID NO:6;
 - (b) the position corresponding to the alanine at amino acid 164 of SEQ ID NO:6;
 - (c) the position corresponding to the glycine at amino acid 165 of SEQ ID NO:6;
 - (d) the position corresponding to the tyrosine at amino acid 370 of SEQ ID NO:6;
 - (e) the position corresponding to the cysteine at amino acid 159 of SEQ ID NO:6;
 - (f) the position corresponding to the isoleucine at amino acid 419 of SEQ ID NO:6;
 - (g) the position corresponding to the alanine at amino acid 175 of SEQ ID NO:6;
- and

- (h) the position corresponding to the leucine at amino acid 337 of SEQ ID NO:6.

More preferred is a DNA molecule wherein the second amino acid substitution occurs at the position corresponding to the serine at amino acid 305 of SEQ ID NO:2 and the first amino acid substitution occurs at a position selected from the group consisting of:

- (a) the position corresponding to the alanine at amino acid 164 of SEQ ID NO:6;
- and

- (b) the position corresponding to the tyrosine at amino acid 370 of SEQ ID NO:6.

Particularly preferred is a DNA molecule wherein the serine occurring at the position corresponding to amino acid 305 of SEQ ID NO:2 is replaced with leucine.

More preferred is a DNA molecule wherein the second amino acid substitution occurs at the position corresponding to the threonine at amino acid 249 of SEQ ID NO:2

and the first amino acid substitution occurs at a position selected from the group consisting of:

- (a) the position corresponding to the alanine at amino acid 164 of SEQ ID NO:6;

and

- (b) the position corresponding to the tyrosine at amino acid 370 of SEQ ID NO:6.

Particularly preferred is a DNA wherein the threonine occurring at the position corresponding to amino acid 249 of SEQ ID NO:2 is replaced with an amino acid selected from the group consisting of isoleucine and alanine.

More preferred is a DNA molecule wherein the second amino acid substitution occurs at the position corresponding to the proline at amino acid 118 of SEQ ID NO:2 and the first amino acid substitution occurs at a position selected from the group consisting of:

- (a) the position corresponding to the alanine at amino acid 164 of SEQ ID NO:6;

and

- (b) the position corresponding to the tyrosine at amino acid 370 of SEQ ID NO:6.

Particularly preferred is a DNA molecule wherein the proline occurring at the position corresponding to amino acid 118 of SEQ ID NO:2 is replaced with a leucine.

More preferred is a DNA molecule wherein the second amino acid substitution occurs at the position corresponding to the asparagine at amino acid 425 of SEQ ID NO:2 and the first amino acid substitution occurs at a position selected from the group consisting of:

- (a) the position corresponding to the alanine at amino acid 164 of SEQ ID NO:6;

and

- (b) the position corresponding to the tyrosine at amino acid 370 of SEQ ID NO:6.

Particularly preferred is a DNA molecule wherein the asparagine occurring at the position corresponding to amino acid 425 of SEQ ID NO:2 is replaced with a serine.

More preferred is a DNA molecule wherein the second amino acid substitution occurs the position corresponding to the tyrosine at amino acid 498 of SEQ ID NO:2 and the first amino acid substitution occurs at a position selected from the group consisting of:

- (a) the position corresponding to the alanine at amino acid 164 of SEQ ID NO:6;

and

- (b) the position corresponding to the tyrosine at amino acid 370 of SEQ ID NO:6.

Particularly preferred is a DNA molecule wherein the tyrosine occurring at the position corresponding to amino acid 498 of SEQ ID NO:2 is replaced with a cysteine.

More preferred is a DNA molecule wherein the tyrosine occurring at the position corresponding to amino acid 370 of SEQ ID NO:6 is replaced with an amino acid selected from the group consisting of cysteine, isoleucine, leucine, threonine, valine and methionine.

Particularly preferred is a DNA molecule wherein the tyrosine occurring at the position corresponding to amino acid 370 of SEQ ID NO:6 is replaced with an amino acid selected from the group consisting of cysteine, isoleucine, leucine, threonine and methionine.

More preferred is a DNA molecule wherein the alanine occurring at the position corresponding to residue 164 of SEQ ID NO:6 is replaced with an amino acid selected from the group consisting of valine, threonine, leucine, cysteine and tyrosine.

More preferred is a DNA molecule wherein the glycine occurring at the position corresponding to residue 165 of SEQ ID NO:6 is replaced with an amino acid selected from the group consisting of serine and leucine.

Particularly preferred is a DNA molecule wherein the glycine occurring at the position corresponding to residue 165 of SEQ ID NO:6 is replaced with a serine.

Particularly preferred is a DNA molecule wherein the arginine occurring at the position corresponding to residue 88 of SEQ ID NO:6 is replaced with a cysteine or a leucine.

More preferred is a DNA molecule wherein the cysteine occurring at the position corresponding to residue 159 of SEQ ID NO:6 is replaced with an amino acid selected from the group consisting of phenylalanine and lysine.

Particularly preferred is a DNA molecule wherein the cysteine occurring at the position corresponding to residue 159 of SEQ ID NO:6 is replaced with a phenylalanine.

More preferred is a DNA molecule wherein the isoleucine occurring at the position corresponding to residue 419 of SEQ ID NO:6 is replaced with an amino acid selected from the group consisting of threonine, histidine, glycine and asparagine.

Particularly preferred is a DNA molecule wherein the isoleucine occurring at the position corresponding to residue 419 of SEQ ID NO:6 is replaced with a threonine.

Particularly preferred is a DNA molecule wherein the alanine occurring at the position corresponding to residue 175 of SEQ ID NO:6 is replaced with an amino acid selected from the group consisting of threonine and valine.

Particularly preferred is a DNA molecule wherein the leucine occurring at the position corresponding to residue 337 of SEQ ID NO:6 is replaced with a serine.

More preferred is a DNA molecule wherein the valine occurring at the position corresponding to residue 356 of SEQ ID NO:10 is replaced with a leucine.

More preferred is a DNA molecule wherein the serine occurring at the position corresponding to residue 421 of SEQ ID NO:10 is replaced with a proline.

More preferred is a DNA molecule wherein the valine occurring at the position corresponding to residue 502 of SEQ ID NO:10 is replaced with an alanine.

More preferred is a DNA molecule wherein the isoleucine occurring at the position corresponding to residue 466 of SEQ ID NO:10 is replaced with a threonine.

More preferred is a DNA molecule wherein the glycine occurring at the position corresponding to residue 212 of SEQ ID NO:10 is replaced with a serine.

More preferred is a DNA molecule wherein the alanine occurring at the position corresponding to residue 211 of SEQ ID NO:10 is replaced with a valine or threonine.

More preferred is a DNA molecule wherein the proline occurring at the position corresponding to residue 369 of SEQ ID NO:12 is replaced with a serine or a histidine.

More preferred is a DNA molecule wherein the alanine occurring at the position corresponding to residue 226 of SEQ ID NO:12 is replaced with a leucine or threonine.

More preferred is a DNA molecule wherein the tyrosine occurring at the position corresponding to residue 432 of SEQ ID NO:12 is replaced with a leucine or isoleucine.

More preferred is a DNA molecule wherein the valine occurring at the position corresponding to residue 517 of SEQ ID NO:12 is replaced with an alanine.

More preferred is a DNA molecule wherein the tyrosine occurring at the position corresponding to residue 428 of SEQ ID NO:16 is replaced with a cysteine, histidine or arginine.

More preferred is a DNA molecule wherein the proline occurring at the position corresponding to residue 365 of SEQ ID NO:16 is replaced with a serine.

More preferred is a DNA molecule wherein the proline occurring at the position corresponding to residue 449 of SEQ ID NO:18 is replaced with an amino acid selected from the group consisting of leucine, isoleucine, valine and methionine.

The present invention is still further directed to a DNA molecule encoding a modified protoporphyrinogen oxidase (protoph) comprising a plant protoph having a double amino acid substitution, wherein both amino acid substitutions are required for there to be resistance to a protoph inhibitor. Preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protoph) comprising a plant protoph, wherein the plant is selected from the group consisting of maize, wheat, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, sugar cane, and *Arabidopsis*. More preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protoph) comprising a plant protoph, wherein the plant is maize.

Preferred is a DNA molecule having a double amino acid substitution, wherein one amino acid substitution occurs at the position corresponding to the leucine at amino acid 347 of SEQ ID NO:6, and wherein the second amino acid substitution occurs at the position corresponding to the alanine at amino acid 453 of SEQ ID NO:6.

Particularly preferred is a DNA molecule having a double amino acid substitution, wherein a leucine occurring at the position corresponding to amino acid 347 of SEQ ID NO:6 is replaced with a serine, and wherein an alanine occurring at the position corresponding to amino acid 453 of SEQ ID NO:6 is replaced with a threonine.

The present invention is directed to expression cassettes and recombinant vectors comprising the expression cassettes comprising essentially a promoter, but especially a promoter that is active in a plant, operatively linked to a DNA molecule encoding the protoporphyrinogen oxidase (protox) enzyme from a eukaryotic organism according to the invention. The expression cassette according to the invention may in addition further comprise a signal sequence operatively linked to the DNA molecule, wherein the signal sequence is capable of targeting the protein encoded by the DNA molecule into the chloroplast or the mitochondria.

The invention relates to a chimeric gene, which comprises an expression cassette comprising essentially a promoter, but especially a promoter that is active in a plant, operatively linked to a heterologous DNA molecule encoding a protoporphyrinogen oxidase (protox) enzyme from a eukaryotic organism according to the invention. Preferred is a chimeric gene, wherein the DNA molecule encodes an protoporphyrinogen oxidase (protox) enzyme from a plant selected from the group consisting of *Arabidopsis*, sugar cane, soybean, barley, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, turf and forage grasses, millet, forage and rice. More preferred is a chimeric gene, wherein the DNA molecule encodes an protoporphyrinogen oxidase (protox) enzyme from a plant selected from the group consisting of soybean, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, turf grass, and rice. Particularly preferred is a chimeric gene, wherein the DNA molecule encodes an protoporphyrinogen oxidase (protox) enzyme from a plant selected from the group consisting of wheat, soybean, cotton, sugar beet, oilseed rape, rice and sorghum. Most preferred is a chimeric gene, wherein the DNA molecule encodes an protoporphyrinogen oxidase (protox) enzyme from a plant selected from the group consisting of soybean, sugar beet, and wheat.

More preferred is a chimeric gene comprising a promoter active in a plant operatively linked to a heterologous DNA molecule encoding a protoporphyrinogen oxidase (protox) selected from the group consisting of a wheat protox comprising the sequence set forth in

SEQ ID NO:10, a soybean protox comprising the sequence set forth in SEQ ID NO:12, cotton protox comprising the sequence set forth in SEQ ID NO:16, a sugar beet protox comprising the sequence set forth in SEQ ID NO:18, an oilseed rape protox comprising the sequence set forth in SEQ ID NO:20, a rice protox comprising the sequence set forth in SEQ ID NO:22, a sorghum protox comprising the sequence set forth in SEQ ID NO:24, and a sugar cane protox comprising the sequence set forth in SEQ ID NO:37. More preferred is a chimeric gene, wherein the protoporphyrinogen oxidase (protox) is selected from the group consisting of a wheat protox comprising the sequence set forth in SEQ ID NO:10, a soybean protox comprising the sequence set forth in SEQ ID NO:12, and a sugar beet protox comprising the sequence set forth in SEQ ID NO:18.

Particularly preferred is a chimeric gene, wherein the DNA molecule encodes a protein from an *Arabidopsis* species having protox-1 activity or protox-2 activity, preferably wherein the protein comprises the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:4.

Particularly preferred is a chimeric gene, wherein the DNA molecule encodes a protein from maize having protox-1 activity or protox-2 activity, preferably wherein the protein comprises the amino acid sequence set forth in set forth in SEQ ID NO:6 or SEQ ID NO:8.

Particularly preferred is a chimeric gene, wherein the DNA molecule encodes a protein from wheat having protox-1 activity preferably wherein the protein comprises the amino acid sequence set forth in SEQ ID NO:10.

Particularly preferred is a chimeric gene, wherein the DNA molecule encodes a protein from soybean having protox-1 activity, preferably wherein the protein comprises the amino acid sequence set forth in SEQ ID NO:12.

Particularly preferred is a chimeric gene, wherein the DNA molecule encodes a protein from cotton having protox-1 activity, preferably wherein the protein comprises the amino acid sequence set forth in SEQ ID NO:16.

Particularly preferred is a chimeric gene, wherein the DNA molecule encodes a protein from sugar beet having protox-1 activity, preferably wherein the protein comprises the amino acid sequence set forth in SEQ ID NO:18.

Particularly preferred is a chimeric gene, wherein the DNA molecule encodes a protein from oilseed rape having protox-1 activity, preferably wherein the protein comprises the amino acid sequence set forth in SEQ ID NO:20.

Particularly preferred is a chimeric gene, wherein the DNA molecule encodes a protein from rice having protox-1 activity, preferably wherein the protein comprises the amino acid sequence set forth in SEQ ID NO:22.

Particularly preferred is a chimeric gene, wherein the DNA molecule encodes a protein from sorghum having protox-1 activity, preferably wherein the protein comprises the amino acid sequence set forth in SEQ ID NO:24.

Particularly preferred is a chimeric gene, wherein the DNA molecule encodes a protein from sugar cane having protox-1 activity, preferably wherein the protein comprises the amino acid sequence set forth in SEQ ID NO:37.

The invention also embodies a chimeric gene, which comprises an expression cassette comprising essentially a promoter, but especially a promoter that is active in a plant, operatively linked to the DNA molecule encoding an protoporphyrinogen oxidase (protox) enzyme from a eukaryotic organism according to the invention, which is resistant to herbicides at levels that inhibit the corresponding unmodified version of the enzyme. Preferred is a chimeric gene, wherein the DNA molecule encodes an protoporphyrinogen oxidase (protox) enzyme from a plant selected from the group consisting of *Arabidopsis*, sugar cane, soybean, barley, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, turf and forage grasses, millet, forage and rice. More preferred is a chimeric gene, wherein the DNA molecule encodes an protoporphyrinogen oxidase (protox) enzyme from a plant selected from the group consisting of soybean, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, turf grass, and rice. Particularly preferred is a chimeric gene, wherein the DNA molecule encodes an protoporphyrinogen oxidase (protox) enzyme from a plant selected from the group consisting of *Arabidopsis*, soybean, cotton, sugar beet, oilseed rape, maize, wheat, sorghum, and rice.

Encompassed by the present invention is a chimeric gene comprising a promoter that is active in a plant operatively linked to the DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a eukaryotic protox having at least one amino acid modification, wherein the amino acid modification has the property of conferring resistance to a protox inhibitor.

Also encompassed by the present invention is a chimeric gene comprising a promoter that is active in a plant operatively linked to the DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox having a first amino acid substitution and a second amino acid substitution; the first amino acid substitution having the property of conferring resistance to a protox inhibitor; and the second amino acid substitution having the property of enhancing the resistance conferred by the first amino

acid substitution. Preferred is the chimeric gene additionally comprising a signal sequence operatively linked to the DNA molecule, wherein the signal sequence is capable of targeting the protein encoded by the DNA molecule into the chloroplast or in the mitochondria.

The chimeric gene according to the invention may in addition further comprise a signal sequence operatively linked to the DNA molecule, wherein the signal sequence is capable of targeting the protein encoded by the DNA molecule into the chloroplast. The chimeric gene according to the invention may in addition further comprise a signal sequence operatively linked to the DNA molecule, wherein the signal sequence is capable of targeting the protein encoded by the DNA molecule into the mitochondria.

Also encompassed by the present invention is any of the DNA sequences mentioned herein before, which is stably integrated into a host genome.

The invention further relates to a recombinant DNA molecule comprising a plant protoporphyrinogen oxidase (protox) or a functionally equivalent derivative thereof.

The invention further relates to a recombinant DNA vector comprising the recombinant DNA molecule of the invention.

A further object of the invention is a recombinant vector comprising the chimeric gene according to the invention, wherein the vector is capable of being stably transformed into a host cell.

A further object of the invention is a recombinant vector comprising the chimeric gene according to the invention, wherein the vector is capable of being stably transformed into a plant, plant seeds, plant tissue or plant cell. Preferred is a recombinant vector comprising the chimeric gene according to the invention, wherein the vector is capable of being stably transformed into a plant. The plant, plant seeds, plant tissue or plant cell stably transformed with the vector is capable of expressing the DNA molecule encoding a protoporphyrinogen oxidase (protox). Preferred is a recombinant vector, wherein the plant, plant seeds, plant tissue or plant cell stably transformed with the the vector is capable of expressing the DNA molecule encoding a protoporphyrinogen oxidase (protox) from a plant that is resistant to herbicides at levels that inhibit the corresponding unmodified version of the enzyme.

Preferred is a recombinant vector comprising the chimeric gene comprising a promoter active in a plant operatively linked to a heterologous DNA molecule encoding a protoporphyrinogen oxidase (protox) selected from the group consisting of a wheat protox comprising the sequence set forth in SEQ ID NO:10, a soybean protox comprising the sequence set forth in SEQ ID NO:12, cotton protox comprising the sequence set forth in SEQ ID NO:16, a sugar beet protox comprising the sequence set forth in SEQ ID NO:18, an

oilseed rape protox comprising the sequence set forth in SEQ ID NO:20, a rice protox comprising the sequence set forth in SEQ ID NO:22, a sorghum protox comprising the sequence set forth in SEQ ID NO:24, and a sugar cane protox comprising the sequence set forth in SEQ ID NO:37, wherein the vector is capable of being stably transformed into a host cell.

Also preferred is recombinant vector comprising the chimeric gene comprising a promoter that is active in a plant operatively linked to the DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox having a first amino acid substitution and a second amino acid substitution; the first amino acid substitution having the property of conferring resistance to a protox inhibitor; and the second amino acid substitution having the property of enhancing the resistance conferred by the first amino acid substitution, wherein the vector is capable of being stably transformed into a plant cell.

Also encompassed by the present invention is a host cell stably transformed with the vector according to the invention, wherein the host cell is capable of expressing the DNA molecule. Preferred is a host cell selected from the group consisting of a plant cell, a bacterial cell, a yeast cell, and an insect cell.

The present invention is further directed to plants and the progeny thereof, plant tissue and plant seeds tolerant to herbicides that inhibit the naturally occurring protox activity in these plants, wherein the tolerance is conferred by a gene expressing a modified inhibitor-resistant protox enzyme as taught herein. Representative plants include any plants to which these herbicides may be applied for their normally intended purpose. Preferred are agronomically important crops, i.e., angiosperms and gymnosperms such as *Arabidopsis*, sugar cane, soybean, barley, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, tomato, potato, turf and forage grasses, millet, forage, and rice and the like. More preferred are agronomically important crops, i.e., angiosperms and gymnosperms such as *Arabidopsis*, cotton, soybean, oilseed rape, sugar beet, maize, rice, wheat, barley, oats, rye, sorghum, millet, turf, forage, turf grasses. Particularly preferred are agronomically important crops, i.e., angiosperms and gymnosperms such as *Arabidopsis*, soybean, cotton, sugar beet, oilseed rape, maize, wheat, sorghum, and rice.

Preferred is a plant comprising the DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox having a first amino acid substitution and a second amino acid substitution; the first amino acid substitution having the property of conferring resistance to a protox inhibitor; and the second amino acid substitution having the property of enhancing the resistance conferred by the first amino acid substitution, wherein the DNA molecule is expressed in the plant and confers upon the

plant tolerance to a herbicide in amounts that inhibit naturally occurring protox activity. Preferred is a plant, wherein the DNA molecule replaces a corresponding naturally occurring protox coding sequence. Comprised by the present invention is a plant and the progeny thereof comprising the chimeric gene according to the invention, wherein the chimeric gene confers upon the plant tolerance to a herbicide in amounts that inhibit naturally occurring protox activity.

Encompassed by the present invention are transgenic plant tissue, including plants and the progeny thereof, seeds, and cultured tissue, stably transformed with at least one chimeric gene according to the invention. Preferred is transgenic plant tissue, including plants, seeds, and cultured tissue, stably transformed with at least one chimeric gene that comprises an expression cassette comprising essentially a promoter, but especially a promoter that is active in a plant, operatively linked to the DNA molecule encoding an protoporphyrinogen oxidase (protox) enzyme that is resistant to herbicides at levels that inhibit the corresponding unmodified version of the enzyme in the plant tissue.

The present invention is further directed to plants, plant tissue, plant seeds, and plant cells tolerant to herbicides that inhibit the naturally occurring protox activity in these plants, wherein the tolerance is conferred by increasing expression of wild-type herbicide-sensitive protox. This results in a level of a protox enzyme in the plant cell at least sufficient to overcome growth inhibition caused by the herbicide. The level of expressed enzyme generally is at least two times, preferably at least five times, and more preferably at least ten times the natively expressed amount. Increased expression may be due to multiple copies of a wild-type protox gene; multiple occurrences of the coding sequence within the gene (*i.e.* gene amplification) or a mutation in the non-coding, regulatory sequence of the endogenous gene in the plant cell. Plants having such altered gene activity can be obtained by direct selection in plants by methods known in the art (see, *e.g.* U.S. Patent No. 5,162,602, and U.S. Patent No. 4,761,373, and references cited therein). These plants also may be obtained by genetic engineering techniques known in the art. Increased expression of a herbicide-sensitive protox gene can also be accomplished by stably transforming a plant cell with a recombinant or chimeric DNA molecule comprising a promoter capable of driving expression of an associated structural gene in a plant cell operatively linked to a homologous or heterologous structural gene encoding the protox enzyme.

The recombinant DNA molecules of the invention can be introduced into the plant cell in a number of art-recognized ways. Those skilled in the art will appreciate that the choice of method might depend on the type of plant, *i.e.* monocot or dicot, targeted for

transformation. Suitable methods of transforming plant cells include microinjection (Crossway *et al.*, *BioTechniques* 4:320-334 (1986)), electroporation (Riggs *et al.*, *Proc. Natl. Acad. Sci. USA* 83:5602-5606 (1986)), *Agrobacterium* mediated transformation (Hinchee *et al.*, *Biotechnology* 6:915-921 (1988)), direct gene transfer (Paszkowski *et al.*, *EMBO J.* 3:2717-2722 (1984)), ballistic particle acceleration using devices available from Agracetus, Inc., Madison, Wisconsin and Dupont, Inc., Wilmington, Delaware (*see, for example*, Sanford *et al.*, U.S. Patent 4,945,050; and McCabe *et al.*, *Biotechnology* 6:923-926 (1988)), protoplast transformation/regeneration methods (*see* U.S. Patent No. 5,350,689 issued Sept. 27, 1994 to Ciba-Geigy Corp.), and pollen transformation (*see* U.S. Patent No. 5,629,183). Also *see*, Weissinger *et al.*, *Annual Rev. Genet.* 22:421-477 (1988); Sanford *et al.*, *Particulate Science and Technology* 5:27-37 (1987)(onion); Christou *et al.*, *Plant Physiol.* 87:671-674 (1988)(soybean); McCabe *et al.*, *Bio/Technology* 6:923-926 (1988)(soybean); Datta *et al.*, *Bio/Technology* 8:736-740 (1990)(rice); Klein *et al.*, *Proc. Natl. Acad. Sci. USA*, 85:4305-4309 (1988)(maize); Klein *et al.*, *Bio/Technology* 6:559-563 (1988)(maize); Klein *et al.*, *Plant Physiol.* 91:440-444 (1988)(maize); Fromm *et al.*, *Bio/Technology* 8:833-839 (1990); Gordon-Kamm *et al.*, *Plant Cell* 2:603-618 (1990)(maize); and U.S. Patent Nos. 5,591,616 and 5,679,558 (rice).

Comprised within the scope of the present invention are transgenic plants, in particular transgenic fertile plants transformed by means of the aforescribed processes and their asexual and/or sexual progeny, which still are resistant or at least tolerant to inhibition by a herbicide at levels that normally are inhibitory to the naturally occurring protox activity in the plant. Progeny plants also include plants with a different genetic background than the parent plant, which plants result from a backcrossing program and still comprise in their genome the herbicide resistance trait according to the invention. Very especially preferred are hybrid plants that are resistant or at least tolerant to inhibition by a herbicide at levels that normally are inhibitory to the naturally occurring protox activity in the plant.

The transgenic plant according to the invention may be a dicotyledonous or a monocotyledonous plant. Preferred are monocotyledonous plants of the *Graminaceae* family involving *Lolium*, *Zea*, *Triticum*, *Triticale*, *Sorghum*, *Saccharum*, *Bromus*, *Oryzae*, *Avena*, *Hordeum*, *Secale* and *Setaria* plants. More preferred are transgenic maize, wheat, barley, sorghum, rye, oats, sugar cane, turf and forage grasses, millet and rice. Especially preferred are maize, wheat, sorghum, rye, oats, turf grasses and rice.

Among the dicotyledonous plants *Arabidopsis*, soybean, cotton, sugar beet, oilseed rape, tobacco, tomato, potato, and sunflower are more preferred herein. Especially preferred are soybean, cotton, tobacco, sugar beet, tomato, potato, and oilseed rape.

The expression 'progeny' is understood to embrace both, "asexually" and "sexually" generated progeny of transgenic plants. This definition is also meant to include all mutants and variants obtainable by means of known processes, such as for example cell fusion or mutant selection and that still exhibit the characteristic properties of the initial transformed plant, together with all crossing and fusion products of the transformed plant material. This also includes progeny plants that result from a backcrossing program, as long as the progeny plants still contain the herbicide resistant trait according to the invention.

Another object of the invention concerns the proliferation material of transgenic plants. The proliferation material of transgenic plants is defined relative to the invention as any plant material that may be propagated sexually or asexually *in vivo* or *in vitro*. Particularly preferred within the scope of the present invention are protoplasts, cells, calli, tissues, organs, seeds, embryos, pollen, egg cells, zygotes, together with any other propagating material obtained from transgenic plants.

Parts of plants, such as for example flowers, stems, fruits, leaves, roots originating in transgenic plants or their progeny previously transformed by means of the process of the invention and therefore consisting at least in part of transgenic cells, are also an object of the present invention.

A further object of the invention is a method of producing plants, protoplasts, cells, calli, tissues, organs, seeds, embryos, pollen, egg cells, zygotes, together with any other propagating material, parts of plants, such as for example flowers, stems, fruits, leaves, roots originating in transgenic plants or their progeny previously transformed by means of the process of the invention, which therefore produce an inhibitor resistant form of a plant protox enzyme by transforming the plant, plant parts with the DNA according to the invention. Preferred is a method of producing a host cell comprising an isolated DNA molecule encoding a protein from a eukaryote having protoporphyrinogen oxidase (protox) activity comprising transforming the host cell with a recombinant vector molecule according to the invention. Further preferred is a method of producing a plant cell comprising an isolated DNA molecule encoding a protein from a eukaryote having protoporphyrinogen oxidase (protox) activity comprising transforming the plant cell with a recombinant vector molecule according to the invention. Preferred is a method of producing transgenic progeny of a transgenic parent plant comprising an isolated DNA molecule encoding a protein from a eukaryote having protoporphyrinogen oxidase (protox) activity comprising transforming the

parent plant with a recombinant vector molecule according to the invention and transferring the herbicide tolerant trait to the progeny of the transgenic parent plant involving known plant breeding techniques.

Preferred is a method for the production of plants, plant tissues, plant seeds and plant parts, which produce an inhibitor-resistant form of the plant protox enzyme, wherein the plants, plant tissues, plant seeds and plant parts have been stably transformed with a structural gene encoding the resistant protox enzyme. Particularly preferred is a method for the production of plants, plant tissues, plant seeds and plant parts, wherein the plants, plant tissues, plant seeds and plant parts have been stably transformed with the DNA according to the invention. Especially preferred is a method for the production of the plants, plant tissues, plant seeds and plant parts, which produce an inhibitor-resistant form of the plant protox enzyme, wherein the plants, plant tissues, plant seeds and plant parts have been prepared by direct selection techniques whereby herbicide resistant lines are isolated, characterized and developed.

The genetic properties engineered into the transgenic seeds and plants described above are passed on by sexual reproduction or vegetative growth and can thus be maintained and propagated in progeny plants. Generally the maintenance and propagation make use of known agricultural methods developed to fit specific purposes such as tilling, sowing or harvesting. Specialized processes such as hydroponics or greenhouse technologies can also be applied. As the growing crop is vulnerable to attack and damages caused by insects or infections as well as to competition by weed plants, measures are undertaken to control weeds, plant diseases, insects, nematodes, and other adverse conditions to improve yield. These include mechanical measures such as tillage of the soil or removal of weeds and infected plants, as well as the application of agrochemicals such as herbicides, fungicides, gametocides, nematocides, growth regulants, ripening agents and insecticides.

Use of the advantageous genetic properties of the transgenic plants and seeds according to the invention can further be made in plant breeding that aims at the development of plants with improved properties such as tolerance of pests, herbicide tolerance, or stress tolerance, improved nutritional value, increased yield, or improved structure causing less loss from lodging or shattering. The various breeding steps are characterized by well-defined human intervention such as selecting the lines to be crossed, directing pollination of the parental lines, or selecting appropriate progeny plants. Depending on the desired properties different breeding measures are taken. The relevant techniques are well known in the art and include but are not limited to hybridization,

inbreeding, backcross breeding, multiline breeding, variety blend, interspecific hybridization, aneuploid techniques, etc. Hybridization techniques also include the sterilization of plants to yield male or female sterile plants by mechanical, chemical or biochemical means. Cross pollination of a male sterile plant with pollen of a different line assures that the genome of the male sterile but female fertile plant will uniformly obtain properties of both parental lines. Thus, the transgenic seeds and plants according to the invention can be used for the breeding of improved plant lines that for example increase the effectiveness of conventional methods such as herbicide or pesticide treatment or allow to dispense with the methods due to their modified genetic properties. Alternatively new crops with improved stress tolerance can be obtained that, due to their optimized genetic "equipment", yield harvested product of better quality than products that were not able to tolerate comparable adverse developmental conditions.

In seeds production germination quality and uniformity of seeds are essential product characteristics, whereas germination quality and uniformity of seeds harvested and sold by the farmer is not important. As it is difficult to keep a crop free from other crop and weed seeds, to control seedborne diseases, and to produce seed with good germination, fairly extensive and well-defined seed production practices have been developed by seed producers, who are experienced in the art of growing, conditioning and marketing of pure seed. Thus, it is common practice for the farmer to buy certified seed meeting specific quality standards instead of using seed harvested from his own crop. Propagation material to be used as seeds is customarily treated with a protectant coating comprising herbicides, insecticides, fungicides, bactericides, nematocides, molluscicides or mixtures thereof. Customarily used protectant coatings comprise compounds such as captan, carboxin, thiram (TMTD[®]), methalaxyl (Apron[®]), and pirimiphos-methyl (Actellic[®]). If desired these compounds are formulated together with further carriers, surfactants or application-promoting adjuvants customarily employed in the art of formulation to provide protection against damage caused by bacterial, fungal or animal pests. The protectant coatings may be applied by impregnating propagation material with a liquid formulation or by coating with a combined wet or dry formulation. Other methods of application are also possible such as treatment directed at the buds or the fruit.

It is thus a further object of the present invention to provide plant propagation material for cultivated plants, but especially plant seed that is treated with an seed protectant coating customarily used in seed treatment.

It is a further aspect of the present invention to provide new agricultural methods such as the methods exemplified above, which are characterized by the use of transgenic

plants, transgenic plant material, or transgenic seed according to the present invention. Comprised by the present invention is an agricultural method, wherein a transgenic plant or the progeny thereof is used comprising a chimeric gene according to the invention in an amount sufficient to express herbicide resistant forms of herbicide target proteins in a plant to confer tolerance to the herbicide.

To breed progeny from plants transformed according to the method of the present invention, a method such as that which follows may be used: maize plants produced as described in the examples set forth below are grown in pots in a greenhouse or in soil, as is known in the art, and permitted to flower. Pollen is obtained from the mature tassel and used to pollinate the ears of the same plant, sibling plants, or any desirable maize plant. Similarly, the ear developing on the transformed plant may be pollinated by pollen obtained from the same plant, sibling plants, or any desirable maize plant. Transformed progeny obtained by this method may be distinguished from non-transformed progeny by the presence of the introduced gene(s) and/or accompanying DNA (genotype), or the phenotype conferred. The transformed progeny may similarly be selfed or crossed to other plants, as is normally done with any plant carrying a desirable trait. Similarly, tobacco or other transformed plants produced by this method may be selfed or crossed as is known in the art in order to produce progeny with desired characteristics. Similarly, other transgenic organisms produced by a combination of the methods known in the art and this invention may be bred as is known in the art in order to produce progeny with desired characteristics.

The modified inhibitor-resistant protox enzymes of the invention have at least one amino acid substitution, addition or deletion relative to their naturally occurring counterpart (i.e. inhibitor-sensitive forms that occur naturally in a plant without being manipulated, either directly *via* recombinant DNA methodology or indirectly *via* selective breeding, etc., by man). Amino acid positions that may be modified to yield an inhibitor-resistant form of the protox enzyme, or enhance inhibitor resistance, are indicated in bold type in Table 1A in the context of plant protox-1 sequences from *Arabidopsis*, maize, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, wheat, and sugar cane. The skilled artisan will appreciate that equivalent changes may be made to any plant protox gene having a structure sufficiently similar to the protox enzyme sequences shown herein to allow alignment and identification of those amino acids that are modified according to the invention to generate inhibitor-resistant forms of the enzyme. Such additional plant protox genes may be obtained using standard techniques as described in International application no. PCT/IB95/00452 filed June 8, 1995, published Dec. 21, 1995 as WO 95/34659 whose relevant parts are herein incorporated by reference.

DNA molecules encoding the herbicide resistant protox coding sequences taught herein may be genetically engineered for optimal expression in a crop plant. This may include altering the coding sequence of the resistance allele for optimal expression in the crop species of interest. Methods for modifying coding sequences to achieve optimal expression in a particular crop species are well known (see, *e.g.* Perlak *et al.*, *Proc. Natl. Acad. Sci. USA* 88: 3324 (1991); Koziel *et al.*, *Bio/technol.* 11: 194 (1993)).

Genetically engineering a protox coding sequence for optimal expression may also include operatively linking the appropriate regulatory sequences (i.e. promoter, signal sequence, transcriptional terminators). Examples of promoters capable of functioning in plants or plant cells (i.e., those capable of driving expression of the associated structural genes such as protox in plant cells) include the cauliflower mosaic virus (CaMV) 19S or 35S promoters and CaMV double promoters; nopaline synthase promoters; pathogenesis-related (PR) protein promoters; small subunit of ribulose biphosphate carboxylase (ssuRUBISCO) promoters, heat shock protein promoter from Brassica with reference to EPA 0 559 603 (hsp80 promoter), *Arabidopsis* actin promoter and the SuperMas promoter with reference to WO 95/14098 and the like. Preferred promoters will be those that confer high level constitutive expression or, more preferably, those that confer specific high level expression in the tissues susceptible to damage by the herbicide. Preferred promoters are the rice actin promoter (McElroy *et al.*, *Mol. Gen. Genet.* 231: 150 (1991)), maize ubiquitin promoter (EP 0 342 926; Taylor *et al.*, *Plant Cell Rep.* 12: 491 (1993)), and the PR-1 promoter from tobacco, *Arabidopsis*, or maize (see U.S. Patent No. 5,614,395 to Ryals *et al.*, incorporated by reference herein in its entirety). The promoters themselves may be modified to manipulate promoter strength to increase protox expression, in accordance with art-recognized procedures.

The inventors have also discovered that another preferred promoter for use with the inhibitor-resistant protox coding sequences is the promoter associated with the native protox gene (i.e. the protox promoter; see copending, co-owned U.S. Patent Application No. 08/808,323, entitled "Promoters from Protoporphyrinogen Oxidase Genes", incorporated by reference herein in its entirety.) The promoter sequence from an *Arabidopsis* protox-1 gene is set forth in SEQ ID NO:13, the promoter sequence from a maize protox-1 gene is set forth in SEQ ID NO:14, and the promoter sequence from a sugar beet protox-1 gene is set forth in SEQ ID NO:26.

Since the protox promoter itself is suitable for expression of inhibitor-resistant protox coding sequences, the modifications taught herein may be made directly on the native protox gene present in the plant cell genome without the need to construct a chimeric gene

with heterologous regulatory sequences. Such modifications can be made via directed mutagenesis techniques such as homologous recombination and selected for based on the resulting herbicide-resistance phenotype (see, e.g. Example 10, Pazkowski *et al.*, *EMBO J.* 7: 4021-4026 (1988), and U.S. Patent No. 5,487,992, particularly columns 18-19 and Example 8). An added advantage of this approach is that besides containing the native protox promoter, the resulting modified gene will also include any other regulatory elements, such as signal or transit peptide coding sequences, which are part of the native gene.

In the event of transformation of the nuclear genome, signal or transit peptides may be fused to the protox coding sequence in chimeric DNA constructs of the invention to direct transport of the expressed protox enzyme to the desired site of action. Examples of signal peptides include those natively linked to the plant pathogenesis-related proteins, e.g. PR-1, PR-2, and the like. See, e.g., Payne *et al.*, *Plant Mol. Biol.* 11:89-94 (1988). Examples of transit peptides include the chloroplast transit peptides such as those described in Von Heijne *et al.*, *Plant Mol. Biol. Rep.* 9:104-126 (1991); Mazur *et al.*, *Plant Physiol.* 85: 1110 (1987); Vorst *et al.*, *Gene* 65: 59 (1988), and mitochondrial transit peptides such as those described in Boutry *et al.*, *Nature* 328:340-342 (1987). Chloroplast and mitochondrial transit peptides are contemplated to be particularly useful with the present invention as protox enzymatic activity typically occurs within the mitochondria and chloroplast. Most preferred for use are chloroplast transit peptides, as inhibition of the protox enzymatic activity in the chloroplasts is contemplated to be the primary basis for the action of protox-inhibiting herbicides (Witkowski and Halling, *Plant Physiol.* 87: 632 (1988); Lehnen *et al.*, *Pestic. Biochem. Physiol.* 37: 239 (1990); Duke *et al.*, *Weed Sci.* 39: 465 (1991)). Also included are sequences that result in localization of the encoded protein to various cellular compartments such as the vacuole. See, for example, Neuhaus *et al.*, *Proc. Natl. Acad. Sci. USA* 88: 10362-10366 (1991) and Chrispeels, *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 42: 21-53 (1991). The relevant disclosures of these publications are incorporated herein by reference in their entirety.

Chimeric genes of the invention may contain multiple copies of a promoter or multiple copies of the protox structural genes. In addition, the construct(s) may include coding sequences for markers and coding sequences for other peptides such as signal or transit peptides, each in proper reading frame with the other functional elements in the DNA molecule. The preparation of such constructs are within the ordinary level of skill in the art.

Useful markers include peptides providing herbicide, antibiotic or drug resistance, such as, for example, resistance to hygromycin, kanamycin, G418, gentamycin, lincomycin, methotrexate, glyphosate, phosphinothricin, or the like. These markers can be used to

select cells transformed with the chimeric DNA constructs of the invention from untransformed cells. Other useful markers are peptidic enzymes that can be easily detected by a visible reaction, for example a color reaction, for example luciferase, β -glucuronidase, or β -galactosidase.

The method of positive selection of genetically transformed cells into which a desired nucleotide sequence can be incorporated by providing the transformed cells with a selective advantage is herein incorporated by reference as WO 94/20627.

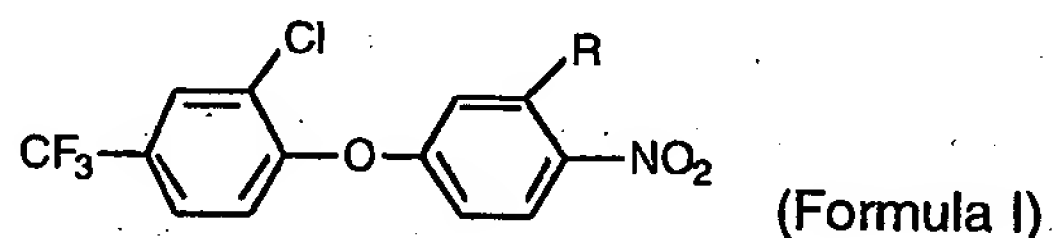
Where a herbicide resistant protox allele is obtained via directed mutation of the native gene in a crop plant or plant cell culture from which a crop plant can be regenerated, it may be moved into commercial varieties using traditional breeding techniques to develop a herbicide tolerant crop without the need for genetically engineering the modified coding sequence and transforming it into the plant. Alternatively, the herbicide resistant gene may be isolated, genetically engineered for optimal expression and then transformed into the desired variety.

Genes encoding altered protox resistant to a protox inhibitor can also be used as selectable markers in plant cell transformation methods. For example, plants, plant tissue or plant cells transformed with a transgene can also be transformed with a gene encoding an altered protox capable of being expressed by the plant. The thus-transformed cells are transferred to medium containing the protox inhibitor wherein only the transformed cells will survive. Protox inhibitors contemplated to be particularly useful as selective agents are the diphenylethers {e.g. acifluorfen, 5-[2-chloro-4-(trifluoromethyl)phenoxy]-2-nitrobenzoic acid; its methyl ester; or oxyfluorfen, 2-chloro-1-(3-ethoxy-4-nitrophenoxy)-4-(trifluorobenzene)}, oxidiazoles, (e.g. oxidiazon, 3-[2,4-dichloro-5-(1-methylethoxy)phenyl]-5-(1,1-dimethylethyl)-1,3,4-oxadiazol-2-(3*H*)-one), cyclic imides (e.g. S-23142, *N*-(4-chloro-2-fluoro-5-propargyloxyphenyl)-3,4,5,6-tetrahydrophthalimide; chlorophthalim, *N*-(4-chlorophenyl)-3,4,5,6-tetrahydrophthalimide), phenyl pyrazoles (e.g. TNPP-ethyl, ethyl 2-[1-(2,3,4-trichlorophenyl)-4-nitropyrazolyl-5-oxy]propionate; M&B 39279), pyridine derivatives, such as e.g. pyridyl amides (e.g. LS 82-556), and phenopylate and its *O*-phenylpyrrolidino- and piperidinocarbamate analogs and triazolones, such as bicyclic triazolones as disclosed in the International patent application WO 92/04827; EP 532146).

The method is applicable to any plant cell capable of being transformed with an altered protox-encoding gene, and can be used with any transgene of interest. Expression of the transgene and the protox gene can be driven by the same promoter functional on plant cells, or by separate promoters.

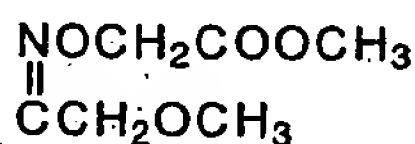
Modified inhibitor-resistant protox enzymes of the present invention are resistant to herbicides that inhibit the naturally occurring protox activity. The herbicides that inhibit protox include many different structural classes of molecules (Duke *et al.*, *Weed Sci.* 39: 465 (1991); Nandihalli *et al.*, *Pesticide Biochem. Physiol.* 43: 193 (1992); Matringe *et al.*, *FEBS Lett.* 245: 35 (1989); Yanase and Andoh, *Pesticide Biochem. Physiol.* 35: 70 (1989)), including the diphenylethers {e.g. acifluorfen, 5-[2-chloro-4-(trifluoromethyl)phenoxy]-2-nitrobenzoic acid; its methyl ester; or oxyfluorfen, 2-chloro-1-(3-ethoxy-4-nitrophenoxy)-4-(trifluorobenzene)}, oxidiazoles (e.g. oxidiazon, 3-[2,4-dichloro-5-(1-methylethoxy)phenyl]-5-(1,1-dimethylethyl)-1,3,4-oxadiazol-2-(3*H*)-one), cyclic imides (e.g. S-23142, *N*-(4-chloro-2-fluoro-5-propargyloxyphenyl)-3,4,5,6-tetrahydrophthalimide; chlorophthalim, *N*-(4-chlorophenyl)-3,4,5,6-tetrahydrophthalimide), phenyl pyrazoles (e.g. TNPP-ethyl, ethyl 2-[1-(2,3,4-trichlorophenyl)-4-nitropyrazolyl-5-oxy]propionate; M&B 39279), pyridine derivatives (e.g. LS 82-556), and phenopylate and its *O*-phenylpyrrolidino- and piperidinocarbamate analogs.

The diphenylethers of particular significance are those having the general formula



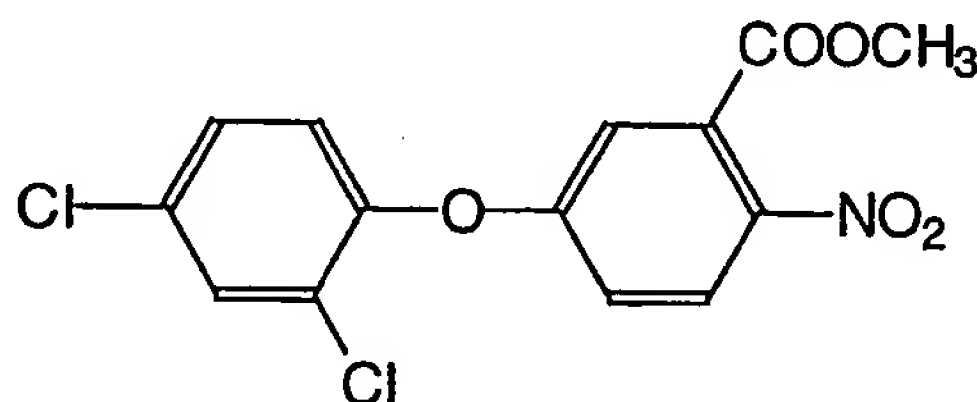
wherein R equals -COONa (Formula II), -CONHSO₂CH₃ (Formula III) or -COOCH₂COOC₂H₅ (Formula IV; see Maigrot *et al.*, *Brighton Crop Protection Conference-Weeds*: 47-51 (1989)).

Additional diphenylethers of interest are those where R equals:



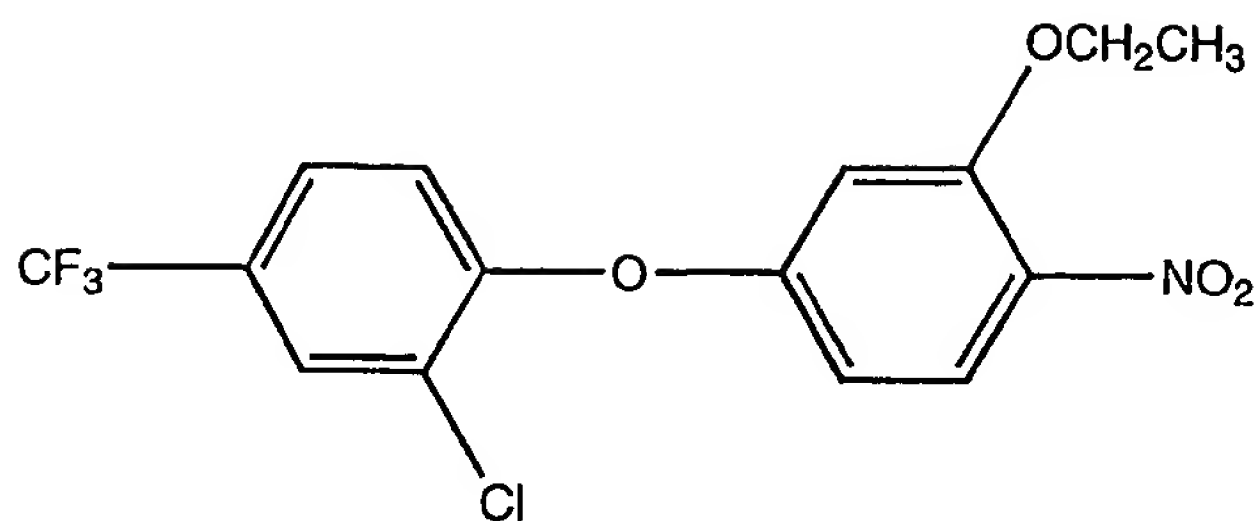
(Formula IVa; see Hayashi *et al.*, *Brighton Crop Protection Conference-Weeds*: 53-58 (1989)).

An additional diphenylether of interest is one having the formula:



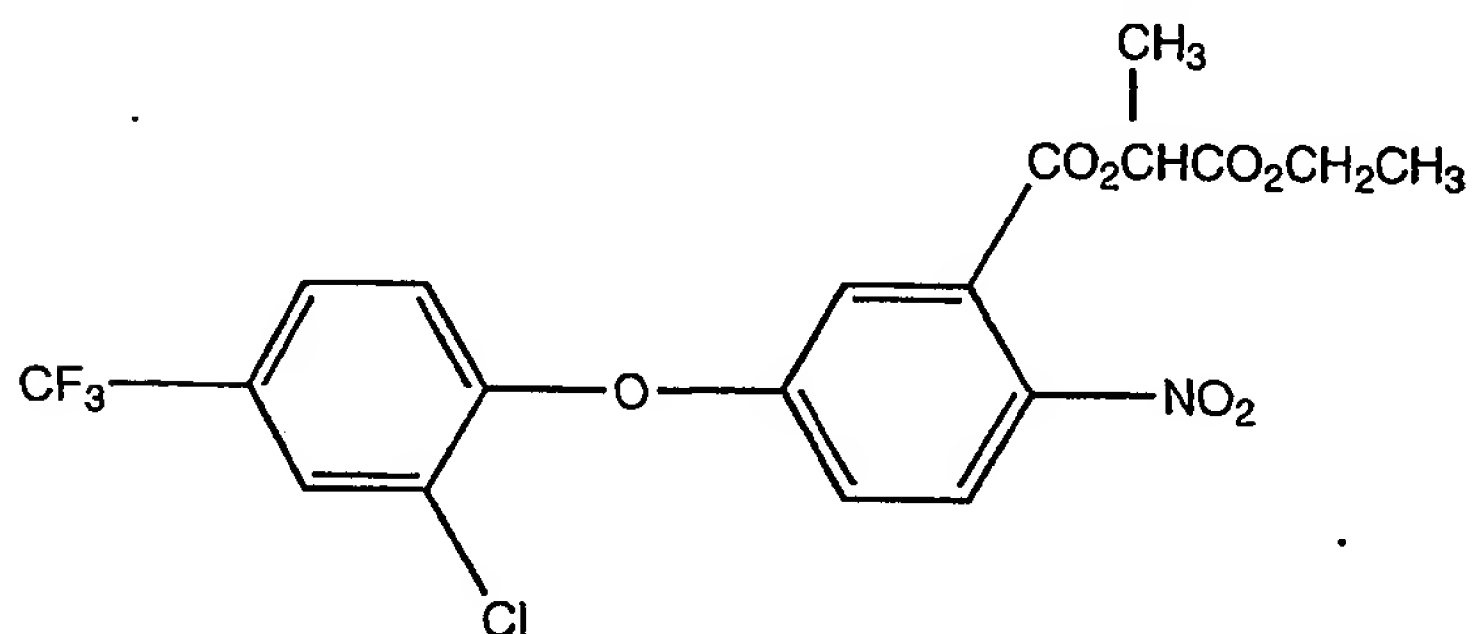
(Formula IVb; bifenox, see Dest *et al.*, *Proc. Northeast Weed Sci. Conf.* 27: 31 (1973)).

A further diphenylether of interest is one having the formula:



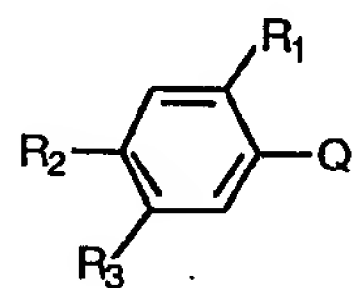
(Formula IVc; oxyfluorfen; see Yih and Swithenbank, *J. Agric. Food Chem.*, 23: 592 (1975))

Yet another diphenylether of interest is one having the formula:



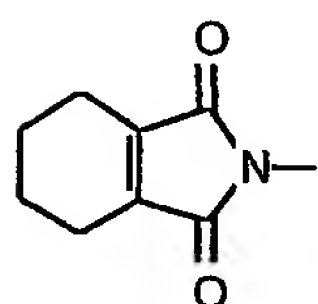
(Formula IVd; lactofen, see page 623 of "The Pesticide Manual", 10th ed., ed. by C. Tomlin, British Crop Protection Council, Surrey (1994))

Also of significance are the class of herbicides known as imides, having the general formula



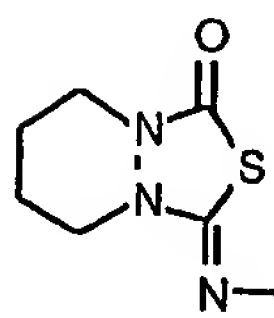
(Formula V)

wherein Q equals



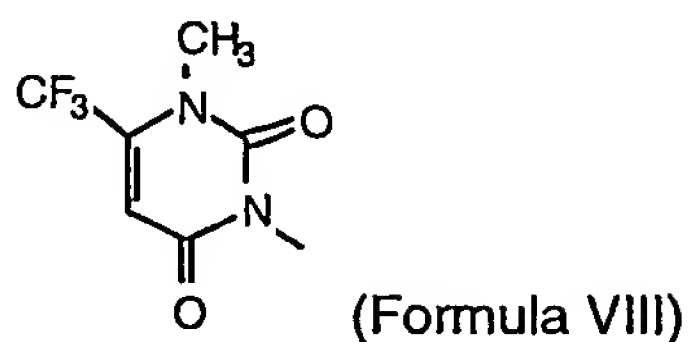
(Formula VI)

or

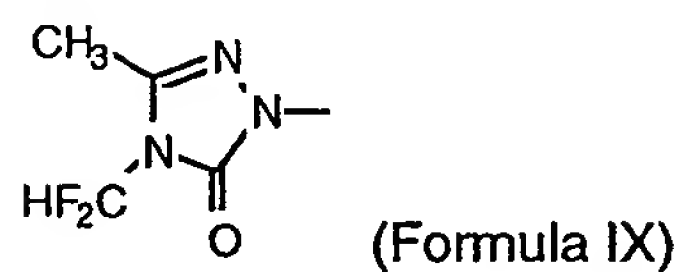


(Formula VII)

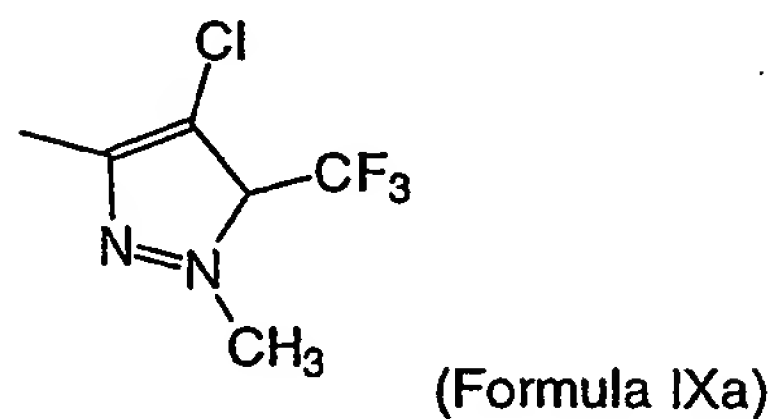
or



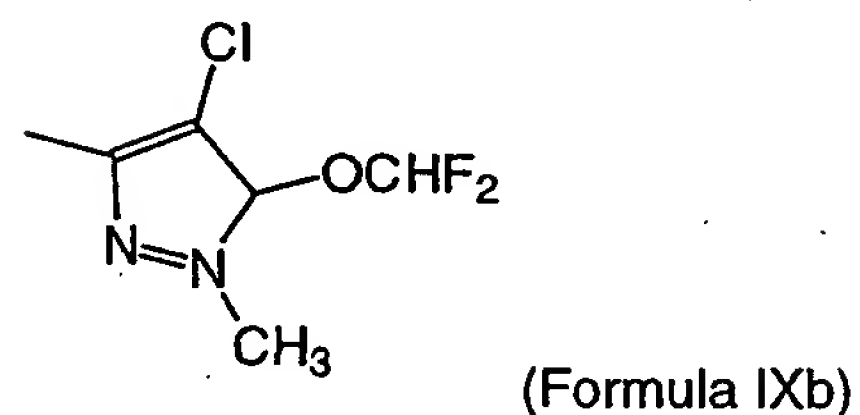
or



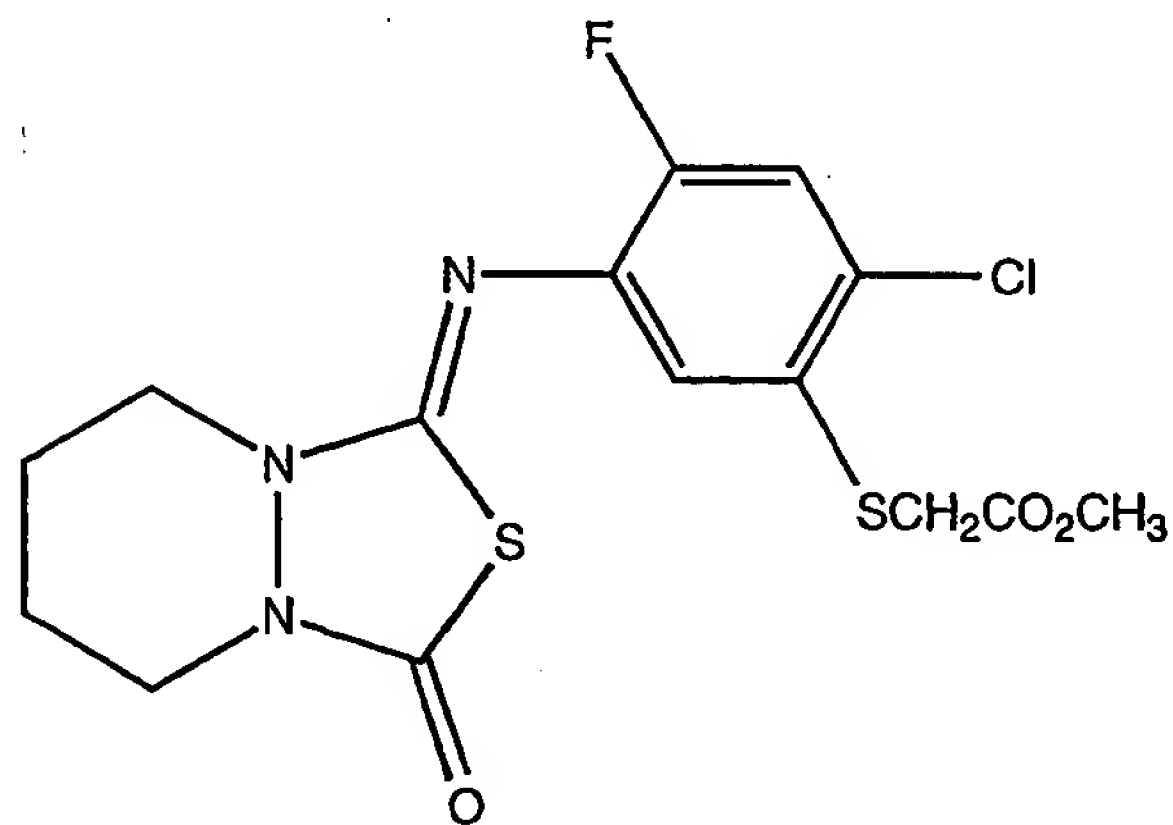
or



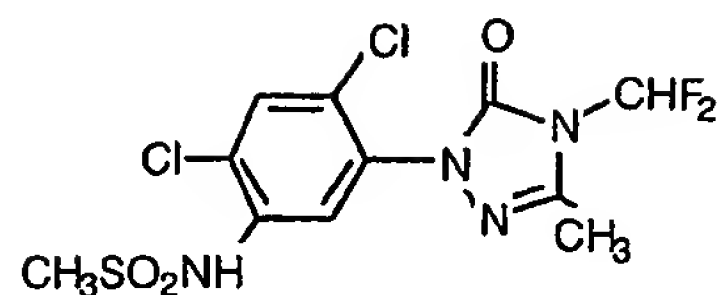
or



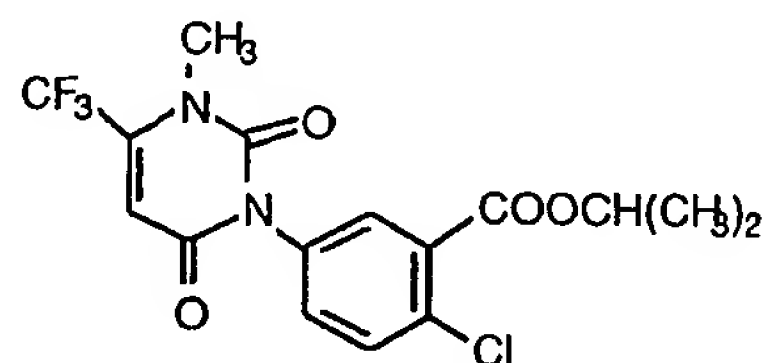
(see Hemper *et al.* (1995) in "Proceedings of the Eighth International Congress of Pesticide Chemistry", Ragdale *et al.*, eds., Amer. Chem. Soc, Washington, D.C., pp.42-48 (1994)); and R_1 equals H, Cl or F, R_2 equals Cl and R_3 is an optimally substituted ether, thioether, ester, amino or alkyl group. Alternatively, R_2 and R_3 together may form a 5 or 6 membered heterocyclic ring. Examples of imide herbicides of particular interest are



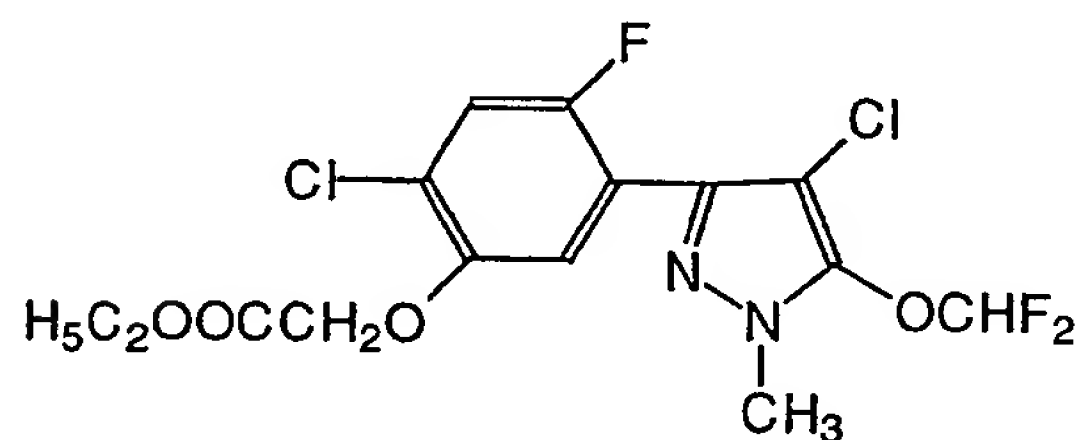
(Formula VIIa; fluthiacet-methyl, see Miyazawa *et al.*, *Brighton Crop Protection Conference-Weeds*, pp. 23-28 (1993))



(Formula X sulfentrazone, see Van Saun *et al.*, *Brighton Crop Protection Conference-Weeds*, pp. 77-82 (1991)).

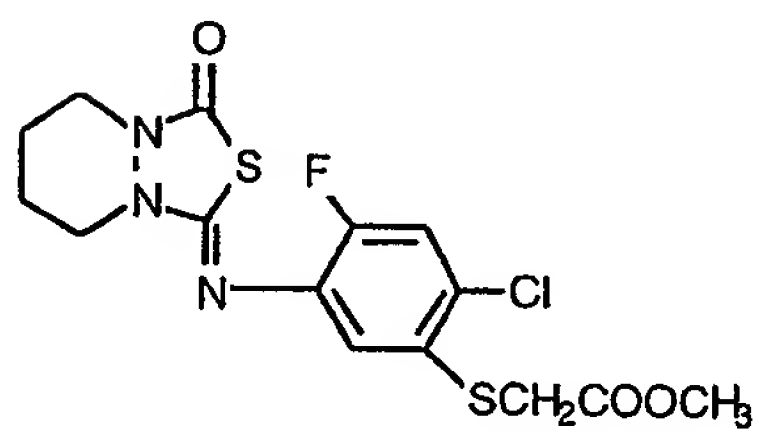


(Formula XI)

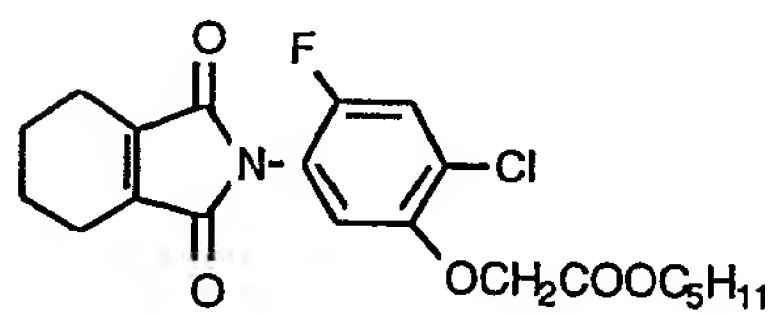


(Formula XII)

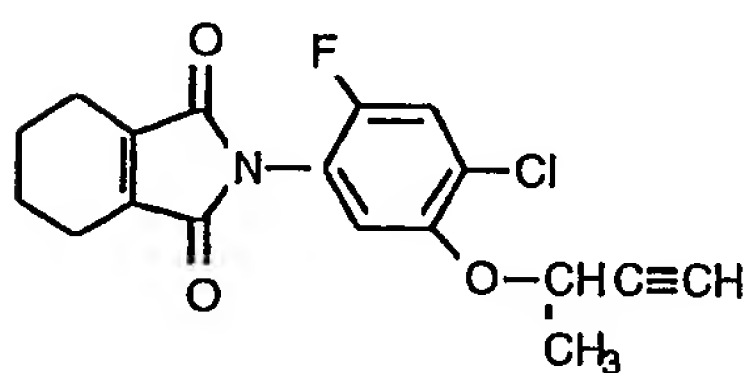
(see Miura *et al.*, *Brighton Crop Protection Conference-Weeds*: 35-40 (1993))



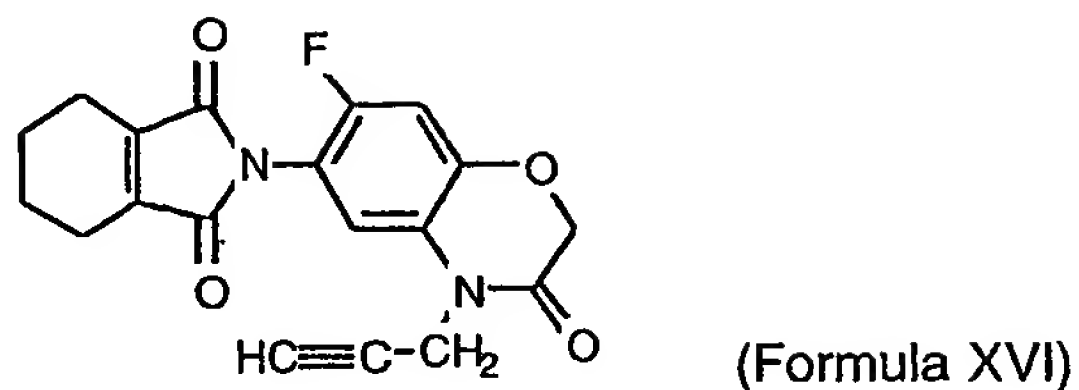
(Formula XIII)



(Formula XIV)

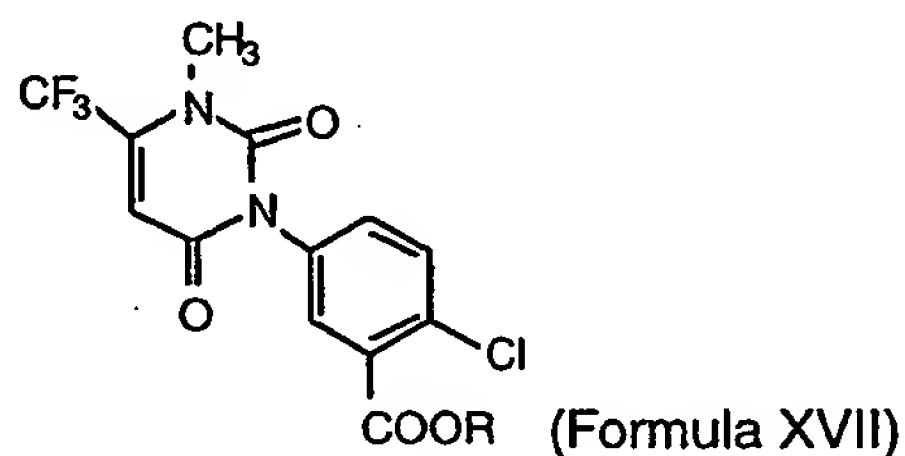


(Formula XV)



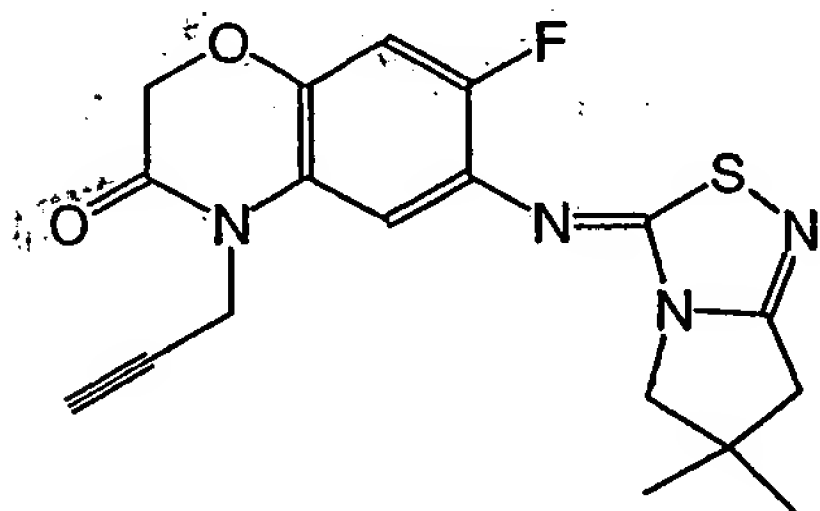
The herbicidal activity of the above compounds is described in the *Proceedings of the 1991 Brighton Crop Protection Conference, Weeds* (British Crop Protection Council) (Formulae X and XVI), *Proceedings of the 1993 Brighton Crop Protection Conference, Weeds* (British Crop Protection Council) (Formulae XII and XIII), U.S. Patent No. 4,746,352 (Formula XI) and *Abstracts of the Weed Science Society of America* vol. 33, pg. 9 (1993)(Formula XIV).

The most preferred imide herbicides are those classified as aryluracils and having the general formula



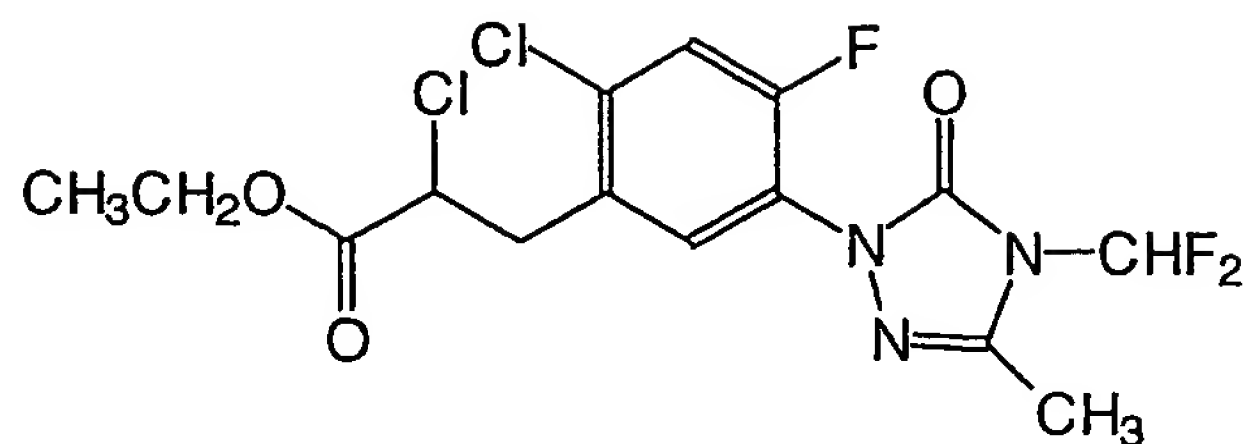
wherein R signifies the group (C₂₋₆-alkenyloxy)carbonyl-C₁₋₄-alkyl, as disclosed in U.S. Patent No. 5,183,492, herein incorporated by reference.

Also of significance are herbicides having the general formula:



(Formula XVIII; thiadiazimin)

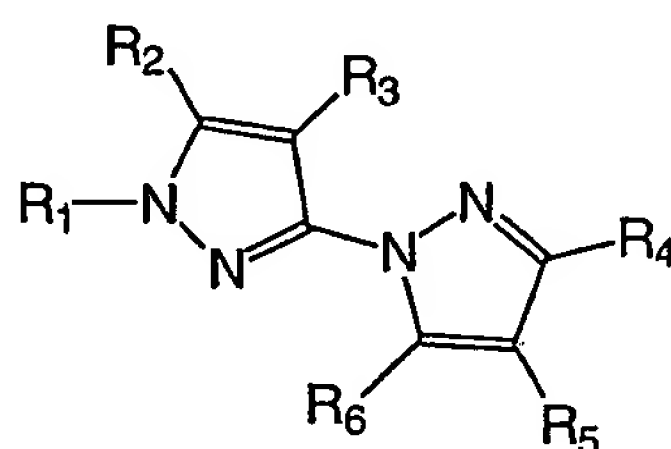
(see Weiler et al., *Brighton Crop Protection Conference-Weeds*, pp. 29-34 (1993));



(Formula XIX; carfentrazone)

(see Van Saun *et al.*, *Brighton Crop Protection Conference-Weeds*: pp. 19-22 (1993));

N-substituted pyrazoles of the general formula:



(Formula XX)

wherein R₁ is C₁-C₄-alkyl, optionally substituted by one or more halogen atoms;

R₂ is hydrogen, or a C₁-C₄-alkoxy, each of which is optionally substituted by one or more halogen atoms, or

R₁ and R₂ together form the group -(CH₂)_n-X-, where X is bound at R₂;

R₃ is hydrogen or halogen,

R₄ is hydrogen or C₁-C₄-alkyl,

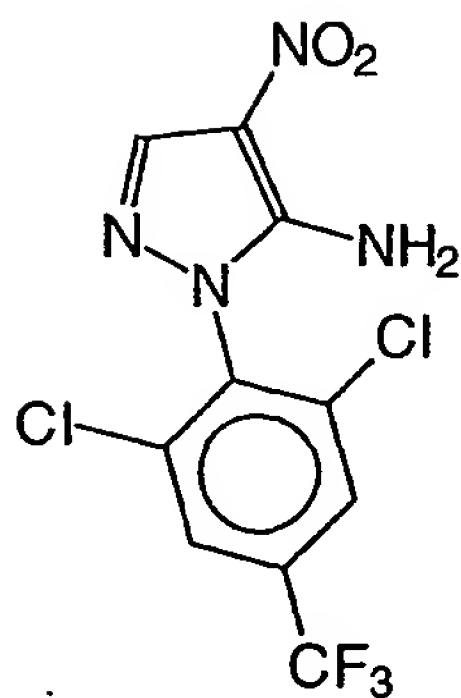
R₅ is hydrogen, nitro, cyano or the group -COOR₆ or -CONR₇R₈, and

R₆ is hydrogen, C₁-C₆-alkyl, C₂-C₆-alkenyl or C₂-C₆-alkynyl;

(see international patent publications WO 94/08999, WO 93/10100, and U. S.

Patent No. 5,405,829 assigned to Schering);

N-phenylpyrazoles, such as:

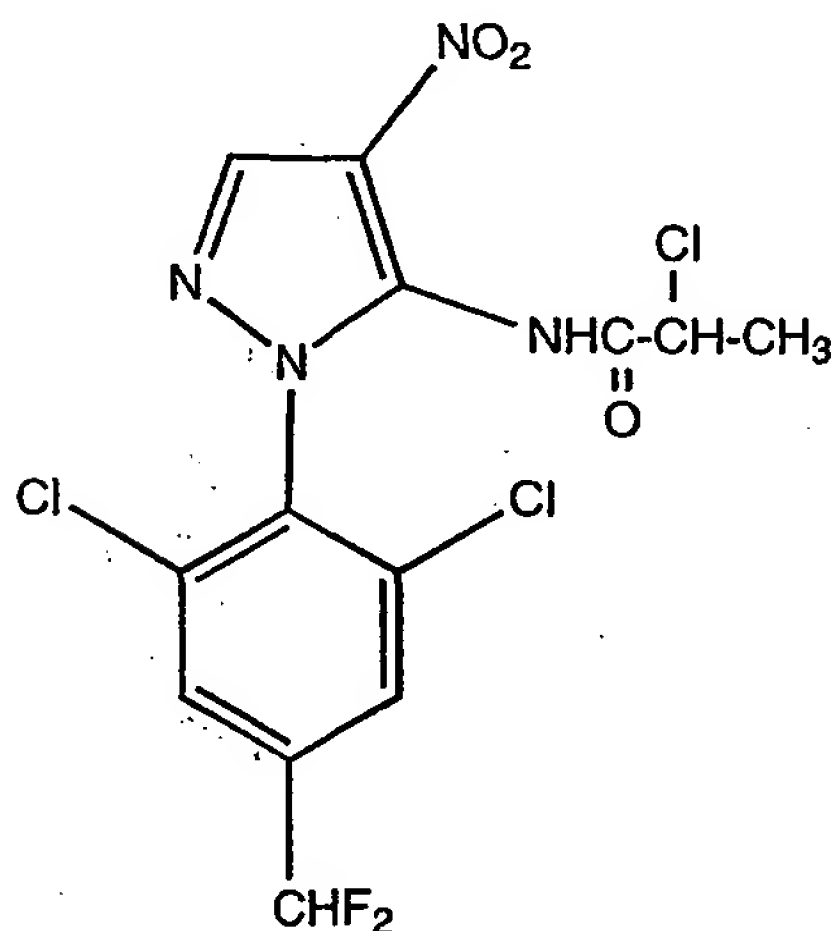


(Formula XXI; nipyraclufen)

(see page 621 of "The Pesticide Manual", 9th ed., ed. by C.R. Worthing, British

Crop Protection Council, Surrey (1991));

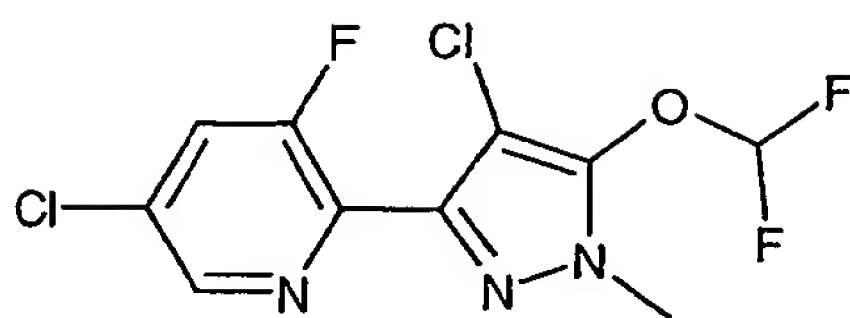
and 3-substituted-2-aryl-4,5,6,7-tetrahydroindazoles (Lyga *et al. Pesticide Sci.* 42:29-36 (1994)).



(Formula XXIa; BAY 11340)

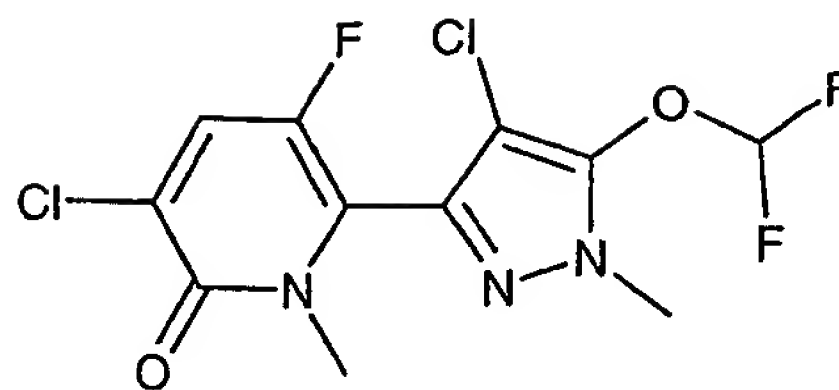
Also of significance are phenylpyrazoles of the type described in WO 96/01254 and WO 97/00246, both of which are hereby incorporated by reference. (Formula XXII).

Also of significance are pyridyl pyrazoles such as the following:



CGA 374410

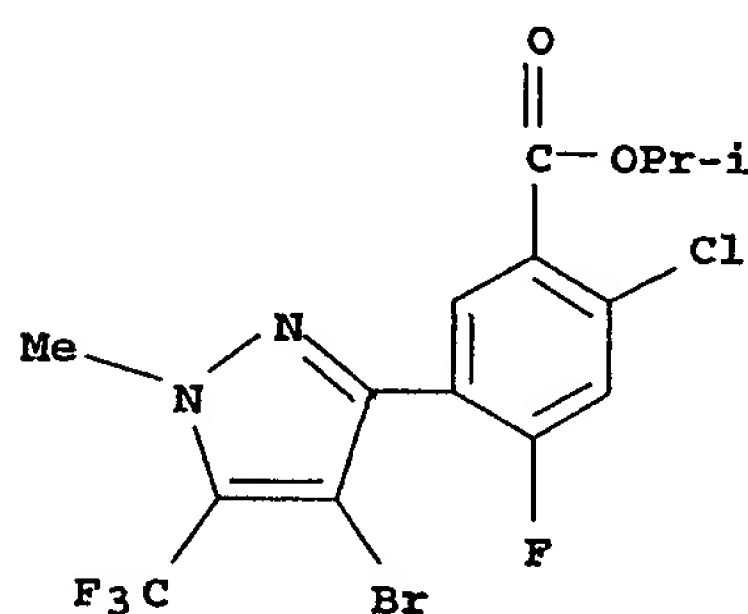
(Formula XXIIIa)



NOA 401954

(Formula XXIIIb)

Also of significance are phenylpyrazoles having the general formula:



(Formula XXIV: Fluazolate)

(see Prosch, S. D. *et al.*, *Brighton Crop Protection Conference-Weeds*: pp. 45-50 (1997), (Vol. 1))

Additional compounds of significance are described in WO98/33927 and US patent 5,856,495, both of which are incorporated herein by reference in their entirety.

Additional compounds of significance are also, for example, Chlornitrofen (4-nitrophenyl 2,4,6-trichlorophenylether), Ethoxyfen (ethyl O-[2-chloro-5-(2-chloro- α,α,α -trifluoro-p-tolyloxy)benzoyl]-L-lactate), Formesafen (5-(2-chloro- α,α,α -trifluoro-p-tolyloxy)-N-methyl-sulfonyl-2-nitro-benzamide), Azafenidin (2-(2,4-dichloro-5-prop-2-ynyloxyphenyl)-5,6,7,8-tetrahydro-1,2,4-triazolo[4,3-a]pyridin-3-(2H)-one), Butafenacil or Fluobutracil (2-[2-chloro-5-(3,6-dihydro-2,6-dioxo-3-methyl-4-trifluoromethyl-1-(2H)-pyrimidinyl)-benzoyloxy]-2-methyl-propionic acid allyl ester), Cinidon-ethyl (2-chloro-3-[2-chloro-5-(cyclohex-1-en-1,2-dicarboximido)phenyl]acrylic acid ethyl ester), Fluthiacet-methyl (methyl[2-chloro-4-fluoro-5-(5,6,7,8-tetrahydro-3-oxo-1H,3H-[1,3,4]thiadiazolo[3,4 a] pyridazin-1-ylideneamino)phenylthio]acetate), Oxadiargyl (5-tert.-butyl-3-[2,4-dichloro-5-(prop-2-ynyloxy)phenyl]-1,3,4-oxadiazol-2(3H)-one), Oxadiazon (5-tert.-butyl-3-(2,4-dichloro-5-isopropoxyphenyl)-

1,3,4-oxa-diazol-2(3H)-one), Pentoxazone (3-(4-chloro-5-cyclopentyloxy-2-fluorophenyl)-5-isopropylidene-1,3-oxazolidine-2,4-dione), Pyraflufen-ethyl (ethyl 2-chloro-5-(4-chloro-5-difluoromethoxy-1-methyl-pyrazol-3-yl)-4-fluorophenoxy acetate), Flumipropyn ((±)-N-[4-chloro-2-fluoro-5-(1-methylprop-2-ynyl)oxy phenyl]cyclohex-1-ene-1,2-dicarboximide), Flupropacil (2-chloro-5-(1,2,3,6-tetrahydro-3-methyl-2,6-dioxo-4-trifluoromethyl-pyrimidin-1-yl)benzoic acid isopropyl ester), Thidiazimin ((Z)-6-(6,7-dihydro-6,6-dimethyl-3H,5H-pyrrolo-[2,1-c]thiadiazol-3-ylideneamino)-7-fluoro-4-(2-propynyl)-2H-1,4-benzoxazin-3-(4H)-one), Benzfendizone (methyl 2-{5-ethyl-2-[4-(1,2,3,6-tetrahydro-3-methyl-2,6-dioxo-4-trifluoromethyl-pyrimidin-1-yl)phenoxy]methyl}phenoxy}propionate), Pyrazogyl or Pyraclonil (1-(3-chloro-4,5,6,7-tetrahydropyrazolo[1,5-a]pyridin-2-yl)-5-(methyl-2-propynylamino)-1H-pyrazole-4-carbonitrile), Profluazole (1-chloro-N-[2-chloro-4-fluoro-5-[(6S, 7aR)-6-fluorotetrahydro-1,3-dioxo-1H-pyrrolo[1,2-c]imidazol-2(3H)-yl]phenyl]methanesulfonamide).

An overview of such compounds is also found in "Peroxidizing Herbicides", ed. by Springer, Berlin, eds. P. Böger & K. Wakabayashi, 1999.

Levels of herbicide that normally are inhibitory to the activity of protox include application rates known in the art, and that depend partly on external factors such as environment, time and method of application. For example, in the case of the imide herbicides represented by Formulae V through IX, and more particularly those represented by Formulae X through XVII, the application rates range from 0.0001 to 10 kg/ha, preferably from 0.005 to 2 kg/ha. This dosage rate or concentration of herbicide may be different, depending on the desired action and particular compound used, and can be determined by methods known in the art.

A further object of the invention is a method for controlling the growth of undesired vegetation that comprises applying to a population of the plant selected from a group consisting of *Arabidopsis*, sugar cane, soybean, barley, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, turf and forage grasses, millet, forage and rice and the like an effective amount of a protox-inhibiting herbicide. Preferred is a method for controlling the growth of undesired vegetation, which comprises applying to a population of the selected from the group consisting of selected from the group consisting of soybean, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, turf grasses and rice an effective amount of a protox-inhibiting herbicide. Particularly preferred is a method for controlling the growth of undesired vegetation, which comprises applying to a population of the selected from the group consisting of *Arabidopsis*, soybean, cotton, sugar beet, oilseed rape, maize, wheat, sorghum, and rice.

III. Plastid Transformation and Expression

The present invention further encompasses a chimeric gene comprising a promoter capable of expression in a plant plastid operatively linked to a DNA molecule of the present invention. A preferred promoter capable of expression in a plant plastid is a promoter isolated from the 5' flanking region upstream of the coding region of a plastid gene, which may come from the same or a different species, and the native product of which is typically found in a majority of plastid types including those present in non-green tissues. Examples of such promoters are promoters of *clpP* genes, such as the tobacco *clpP* gene promoter (WO 97/06250, incorporated herein by reference) and the *Arabidopsis clpP* gene promoter (U.S. Application No. 09/038,878, incorporated herein by reference). Other promoters that are capable of expressing a DNA molecule in plant plastids are promoters recognized by viral RNA polymerases. Preferred promoters of this type are promoters recognized by a single sub-unit RNA polymerase, such as the T7 gene 10 promoter, which is recognized by the bacteriophage T7 DNA-dependent RNA polymerase. Yet another promoter that is capable of expressing a DNA molecule in plant plastids comes from the regulatory region of the plastid 16S ribosomal RNA operon (Harris *et al.*, *Microbiol. Rev.* 58:700-754 (1994), Shinozaki *et al.*, *EMBO J.* 5:2043-2049 (1986), both of which are incorporated herein by reference). The gene encoding the T7 polymerase is preferably transformed into the nuclear genome and the T7 polymerase is targeted to the plastids using a plastid transit peptide. Expression of the DNA molecules in the plastids can be constitutive or can be inducible. These different embodiment are extensively described in WO 98/11235, incorporated herein by reference. The chimeric gene preferably further comprises a 5' untranslated sequence (5' UTR) functional in plant plastids and a plastid gene 3' untranslated sequence (3' UTR) operatively linked to a DNA molecule of the present invention. Preferably, the 3' UTR is a plastid *rps16* gene 3' untranslated sequence. In a further embodiment, the chimeric gene comprises a poly-G tract instead of a 3' untranslated sequence.

The present invention also encompasses a plastid transformation vector comprising the chimeric gene described above and flanking regions for integration into the plastid genome by homologous recombination. The plastid transformation vector may optionally comprise at least one chloroplast origin of replication. The present invention also encompasses a plant plastid transformed with such a plastid transformation vector, wherein the DNA molecule is expressible in the plant plastid. The invention also encompasses a plant or plant cell, including the progeny thereof, comprising this plant plastid. In a preferred

embodiment, the plant is homoplasmic for transgenic plastids. The plants transformed in the present invention may be monocots or dicots. A preferred monocot is maize and a preferred dicot is tobacco. Other preferred dicots are tomato and potato.

In a preferred embodiment, the present invention encompasses a chimeric gene comprising a promoter capable of expression in a plant plastid operatively linked to a DNA molecule isolated from a prokaryote or a eukaryote that encodes a native or modified protox enzyme, such as a DNA molecule that encodes a native or modified wheat, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, or sugar cane protox enzyme. Such a DNA molecule is comprised in a plastid transformation vector as described above and plants homoplasmic for transgenic plastid genomes are produced. Expression in plant plastids of a DNA molecule that encodes a modified protox enzyme preferably confers upon the plant tolerance to a herbicide in amounts that inhibit naturally occurring protox activity.

In a further preferred embodiment, the present invention encompasses a chimeric gene comprising (a) a DNA molecule isolated from a plant, which in its native state encodes a polypeptide that comprises a plastid transit peptide, and a mature enzyme that is natively targeted to a plastid of the plant by the plastid transit peptide, wherein the DNA molecule is modified such that it does not encode a functional plastid transit peptide; and (b) a promoter capable of expressing the DNA molecule in a plastid, wherein the promoter is operatively linked to the DNA molecule. In one preferred embodiment, the transit peptide is mutated and thus does not allow the proper transport of the enzyme encoded by the DNA molecule to the desired cell compartment, such as the plastid. In another preferred embodiment, a portion of the transit peptide coding sequence or the entire transit peptide coding sequence is removed from the DNA molecule, preventing the enzyme from being properly targeted to the desired cell compartment.

The chimeric genes described above are inserted in plastid transformation vectors, and the present invention is therefore also directed to plants having their plastid genome transformed with such vectors, whereby the DNA molecule is expressible in plant plastids. Such plants are preferably homoplasmic for transgenic plastids.

In a preferred embodiment, a DNA molecule described immediately above encodes an enzyme that in its wild-type form is inhibited by a herbicide. In a further preferred embodiment, the DNA molecule encodes an enzyme that in its wild-type form is inhibited by a herbicide, but that comprises at least one amino acid change compared to the wild-type enzyme. Such an amino acid change makes the enzyme resistant to compounds that naturally inhibit the wild-type enzyme. In a further preferred embodiment, the DNA molecule encodes an enzyme having protoporphyrinogen oxidase (protox) activity. In a further

preferred embodiment, the transit peptide is removed from the DNA molecule as further illustrated in Examples 37-42. Plants homoplasmic for transgenic plastids of the invention are resistant to high amounts of herbicides such as Formula XVII that inhibit the naturally occurring protox activity (as further illustrated in Example 44).

In another preferred embodiment, the transit peptide of a DNA molecule encoding a 5-enolpyruvyl-3-phosphoshikimate synthase (EPSP synthase) is mutated or removed. The resulting DNA molecule is fused to a promoter capable of expression in plant plastids and homoplasmic plants harboring such constructs in their plastid genomes are obtained. These plants are resistant to herbicidal compounds that naturally inhibit EPSP synthase, in particular glyphosate. In another preferred embodiment, the transit peptide of a DNA molecule encoding a acetolactate synthase (ALS) is mutated or removed. The resulting DNA molecule is fused to a promoter capable of expression in plant plastids and homoplasmic plants harboring such constructs in their plastid genome are obtained. These plants are resistant to herbicidal compounds that naturally inhibit ALS, in particular sulfonylureas. In another preferred embodiment, the transit peptide of a DNA molecule encoding a acetoxhydroxyacid synthase (AHAS) is mutated or removed. The resulting DNA molecule is fused to a promoter capable of expression in plant plastids and homoplasmic plants harboring such constructs in their plastid genome are obtained. These plants are resistant to herbicidal compounds that naturally inhibit AHAS, in particular, imidazolinone and sulfonamide herbicides. In another preferred embodiment, the transit peptide of a DNA molecule encoding an acetylcoenzyme A carboxylase (ACCase) is mutated or removed. The resulting DNA molecule is fused to a promoter capable of expression in plant plastids and homoplasmic plants harboring such constructs in their plastid genome are obtained. These plants are resistant to herbicidal compounds that naturally inhibit ACCase, in particular cyclohexanedione and arylphenoxypropanoic acid herbicides. In another preferred embodiment, the transit peptide of a DNA molecule encoding a glutamine synthase (GS) is mutated or removed. The resulting DNA molecule is fused to a promoter capable of expression in plant plastids and homoplasmic plants harboring such constructs in their plastid genome are obtained. These plants are resistant to herbicidal compounds that naturally inhibit GS, in particular phosphinothricin and methionine sulfoximine.

The present invention is also further directed to a method of obtaining herbicide-resistant plants by transforming their plastid genome with a chimeric gene comprising (a) a DNA molecule isolated from a plant, which in its native state encodes a polypeptide that comprises a plastid transit peptide, and a mature enzyme that is natively targeted to a

plastid of the plant by the plastid transit peptide, wherein the DNA molecule is modified such that it does not encode a functional plastid transit peptide; and (b) a promoter capable of expressing the DNA molecule in a plastid, wherein the promoter is operatively linked to the DNA molecule. Examples of enzymes that are used in the present invention are cited immediately above, but the applicability of such a method is not limited to the cited examples.

The present invention is still further directed to a novel method for selecting a transplastomic plant cell, comprising the steps of: introducing the above-described chimeric gene into the plastome of a plant cell; expressing the encoded enzyme in the plastids of the plant cell; and selecting a cell that is resistant to a herbicidal compound that naturally inhibits the activity of the enzyme, whereby the resistant cell comprises transformed plastids. In a preferred embodiment, the enzyme is naturally inhibited by a herbicidal compound and the transgenic plant is able to grow on an amount of the herbicidal compound that naturally inhibits the activity of the enzyme. In a further preferred embodiment, the enzyme has protoporphyrinogen oxidase (protoporphyrinogen oxidase) activity and is modified so that it confers resistance to protoporphyrinogen oxidase inhibitors.

A further aspect of the present invention is a novel method for plastid transformation of recalcitrant plants. The methods pioneered for plastid transformation of tobacco and lower plant species rely on non-lethal selection for resistance to antibiotics that preferentially affect the plastid translational apparatus and hence allow photo-heterotrophic transformants to outgrow heterotrophic, non-transformed tissue.

Several factors have likely contributed to the difficulties encountered with plastid transformation of monocots and other dicots. For example, the maize chloroplast 16S ribosomal RNA (rRNA) is naturally resistant to spectinomycin because of the presence of a G at position 1138 in the *Zea mays* 16S rDNA gene (Harris *et al.*, 1994). Thus, utilization of 16S rRNA point mutations that confer spectinomycin and/or streptomycin resistance which have been used successfully as selectable chloroplast markers in *Chlamydomonas* and tobacco (Boynton and Gillham (1993) *In* Wu, R. [Ed.] *Methods in Enzymology* Vol 217. Academic Press, San Diego, pp. 510-536; Svab *et al.* (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87: 8526-8530) is not feasible for maize. Natural spectinomycin and streptomycin resistance in maize also obviates the use of the bacterial *aadA* gene encoding aminoglycoside 3'-adenyltransferase, which results in dominant spectinomycin and streptomycin resistance and allows a 100-fold increase in tobacco chloroplast transformation efficiency (Svab and Maliga (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90: 913-917). Use of kanamycin (the only other antibiotic proven to be useful for chloroplast

transformation) is also problematic due to a large excess (ca. 50:1) of nuclear vs. chloroplast-encoded resistance in tobacco following bombardment of the bacterial *nptII* gene encoding neomycin phosphotransferase (Carrer *et al.* (1993) *Mol. Gen. Genet.* 241: 49-56). This has been shown to result from both a high frequency of spontaneous nuclear resistance mutants as well as integration of *nptII* into the nuclear genome. Since *nptII* is also a highly effective selectable marker for maize nuclear transformation it is reasonable to expect similar background levels to that observed in tobacco. Spontaneous resistance and a significant excess of selectable marker integration by random, illegitimate recombination into the nuclear genome, rather than homologous integration into the chloroplast genome, would make recovery of bona fide chloroplast transformants difficult if not impossible.

A more fundamental reason for the difficulties encountered with plastid transformation in plant species other than tobacco may have to do with the non-photosynthetic nature of many regenerable cultured plant tissues, especially in maize and *Arabidopsis*. Tobacco is an exception in that cultured vegetative tissues are regenerable and contain mature differentiated chloroplasts that are photosynthetically competent in the presence of sucrose. Consequently, the current system for selecting tobacco plastid transformants relies on the faster growth rate of transformed cells that can use both reduced and inorganic carbon sources. Moreover, transformed cells do not suffer the chloroplast membrane damage that results from inhibition of plastid protein synthesis in the light. This expression of selectable markers that act preferentially on photosynthetic cells, driven by promoters that have high activity in differentiated chloroplasts, is unlikely to work in non-green tissues containing proplastids (e.g. dark-grown maize Type I callus, somatic embryos) or amyloplasts/leucoplasts (e.g. *Arabidopsis* root cultures). Plastid transformation in these plants requires a selectable marker that gives strong selection in all plastid types.

A preferred selectable marker for generalized plastid transformation: (1) is active only in the plastid to eliminate nuclear-transformed "escapes"; (2) has a mode of action that does not depend on photosynthetic competence or the presence of fully differentiated chloroplasts; and (3) has a level of resistance that is co-dependent on an adjustable external parameter (e.g. light), rather than being determined solely by the bulk concentration of a selective agent, so that selection pressure can vary during selection to facilitate segregation of the many-thousand plastid genome copies.

In a preferred embodiment, such a selectable marker gene involves the use of a chimeric gene comprising an isolated DNA molecule encoding a plastid-targeted enzyme having in its natural state a plastid transit peptide, wherein the DNA molecule is modified such that the transit peptide either is absent or does not function to target the enzyme to the

plastid, wherein the DNA molecule is operatively linked to a promoter capable of expression in plant plastids. In a preferred embodiment, a DNA molecule of the present invention encodes an enzyme that is naturally inhibited by a herbicide. In another preferred embodiment, the DNA molecule encodes a protoporphyrinogen IX oxidase ("protox"). In a preferred embodiment, the protoporphyrinogen IX oxidase gene is from *Arabidopsis thaliana* and in a more preferred embodiment, the protoporphyrinogen IX oxidase gene is from *Arabidopsis thaliana* and comprises at least one amino acid substitution. Preferably, an amino acid substitution results in tolerance of the enzyme against inhibition by an herbicide which naturally inhibits the activity of the enzyme. Low concentrations of herbicide are thought to kill wildtype plants due to light-sensitive intermediates which build up when the plastid-localized protox enzyme is inhibited. Production of these photosensitizing compounds does not require differentiated chloroplasts or active photosynthesis, which is a key factor for successful plastid transformation of plants whose regenerable cultured tissues are of non-photosynthetic nature.

Another key feature is to have expression of the selectable marker gene in non-green plastids. In a preferred embodiment, the invention encompasses the use of promoters that are capable of expression of operatively linked DNA molecules in plastids of both green and non-green tissue. In particular, one such promoter comes from the regulatory region of the plastid 16S-ribosomal RNA operon. Another candidate is the promoter and 5' UTR from the plastid *clpP* gene. The *clpP* gene product is expressed constitutively in plastids from all plant tissues, including those that do not contain chloroplasts (Shanklin (1995) *Plant Cell* 7: 1713-22).

Other DNA molecules may be co-introduced in plant plastids using the method described above. In a preferred embodiment, a plastid transformation vector of the present invention contains a chimeric gene allowing for selection of transformants as described above and at least one other gene fused to a promoter capable of expression in plant plastids. The other such gene may, for example, confer resistance to insect pests, or to fungal or bacterial pathogens, or may encode one or more value-added traits.

EXAMPLES

The invention will be further described by reference to the following detailed examples. These examples are provided for purposes of illustration only, and are not intended to be limiting unless otherwise specified. Standard recombinant DNA and molecular cloning techniques used here are well known in the art and are described by

Ausubel (ed.), *Current Protocols in Molecular Biology*, John Wiley and Sons, Inc. (1994); T. Maniatis, E. F. Fritsch and J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor laboratory, Cold Spring Harbor, NY (1989); and by T.J. Silhavy, M.L. Berman, and L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984).

Section A. Isolation And Characterization Of Plant Protoporphyrinogen Oxidase (Protox) Genes

Example 1: Isolation of a Wheat Protox-1 cDNA Based on Sequence Homology to a Maize Protox-1 Coding Sequence

Total RNA prepared from *Triticum aestivum* (cv Kanzler) was submitted to Clontech for custom cDNA library construction in the Lambda Uni-Zap vector. Approximately 50,000 pfu of the cDNA library were plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5; see Example 2 of International application no. PCT/IB95/00452, filed June 8, 1995, published Dec. 21, 1995 as WO 95/34659) labeled with ^{32}P -dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO_4 pH 7.0, 1 mM EDTA at 50°C . Wash conditions were 2X SSC, 1% SDS at 50°C . (Church and Gilbert, *Proc. Natl. Acad. Sci. USA* 81: 1991-1995 (1984), hereby incorporated by reference in its entirety.) Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequences of the cDNA inserts were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest wheat protox-1 cDNA obtained from initial screening efforts, designated "wheat protox-1", was 1489-bp in length. Wheat protox-1 lacks coding sequence for the transit peptide plus approximately 126 amino acids of the mature coding sequence based on comparison with the other known plant protox peptide sequences.

A second screen was performed to obtain a longer wheat protox cDNA. For this screen, a *Triticum aestivum* (cv Kanzler) cDNA library was prepared internally using the lambda Uni-Zap vector. Approximately 200,000 pfu of the cDNA library was screened as indicated above, except that the wheat protox-1 cDNA was used as a probe and hybridization and wash conditions were at 65°C instead of 50°C . The longest wheat cDNA obtained from this screening effort, designated "wheat protox-1a", was 1811-bp in length. The nucleotide sequence of this cDNA and the amino acid sequence it encodes are set

forth in SEQ ID NOs:9 and 10, respectively. Based on comparison with the other known plant protox peptide sequences and with corresponding genomic sequence, this cDNA is either full-length or missing only a few transit peptide codons (Table 1A). This wheat protein sequence is 91% identical (95% similar) to the maize protox-1 protein sequence set forth in SEQ ID NO:6.

Wheat protox-1a, in the pBluescript SK vector, was deposited March 19, 1996, as pWDC-13 (NRRL #B21545).

Example 2: Isolation of a Soybean Protox-1 cDNA Based on Sequence Homology to an *Arabidopsis* Protox-1 Coding Sequence

A Lambda Uni-Zap cDNA library prepared from soybean (v Williams 82, epicotyls) was purchased from Stratagene. Approximately 50,000 pfu of the library was plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto Colony/Plaque Screen membranes (NEN Dupont). The plaque lifts were probed with the *Arabidopsis* protox-1 cDNA (SEQ ID NO:1; see Example 1 of International application no. PCT/IB95/00452, filed June 8, 1995, published Dec. 21, 1995 as WO 95/34659)) labeled with ³²P-dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50° C. Wash conditions were 2X SSC, 1% SDS at 50° C. (Church and Gilbert (1984)). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequence of the cDNA inserts was determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest soybean cDNA obtained, designated "soybean protox-1", is full-length based on comparison with the other known plant protox peptide sequences (Table 1A). Soybean protox-1 is 1847-bp in length and encodes a protein of 58.8 kDa. The nucleotide sequence of this cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs:11 and 12, respectively. The soybean protein is 78% identical (87% similar) to the *Arabidopsis* protox-1 protein.

Soybean protox-1, in the pBluescript SK vector, was deposited December 15, 1995 as pWDC-12 (NRRL #B-21516).

Example 3: Isolation of a Cotton Protox-1 cDNA Based on Sequence Homology to a Maize Protox-1 Coding Sequence

A Lambda Uni-Zap cDNA library prepared from *Gossypium hirsutum* L. (72 hr. dark grown cotyledons) was obtained from Dr. Dick Trelease, Dept. of Botany, Arizona State

University (Ni W. and Trelease R.N., *Arch. Biochem. Biophys.* 289: 237-243 (1991)). Approximately 50,000 pfu of the library was plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto Colony/Plaque Screen membranes (NEN Dupont). The plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5) labeled with ^{32}P -dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO_4 pH 7.0, 1 mM EDTA at 50°C . Wash conditions were 2X SSC, 1% SDS at 50°C . (Church and Gilbert (1984)). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequence of the cDNA inserts was determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest cotton cDNA obtained, designated "cotton protox-1", appears to be full-length based on comparison with the other known plant protox peptide sequences (Table 1A). Cotton protox-1 is 1826-bp in length and encodes a protein of 58.2 kDa. The nucleotide sequence of this cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs:13 and 14, respectively. The cotton protein is 77% identical (86% similar) to the maize protox-1 protein.

Cotton protox-1, in the pBluescript SK vector, was deposited July 1, 1996 as pWDC-15 (NRRL #B-21594).

Example 4: Isolation of a Sugar Beet Protox-1 cDNA Based on Sequence Homology to an *Arabidopsis* Protox-1 Coding Sequence

A Lambda-Zap cDNA library prepared from *Beta vulgaris* was obtained from Dr. Philip Rea, Dept. of Botany, Plant Science Institute, Philadelphia, PA (Yongcheol Kim, Eugene J. Kim, and Philip A. Rea, *Plant Physiol.* 106: 375-382 (1994)). Approximately 50,000 pfu of the cDNA library were plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The plaque lifts were probed with the *Arabidopsis* protox-1 cDNA (SEQ ID NO:1) labeled with ^{32}P -dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO_4 pH 7.0, 1 mM EDTA at 50°C . Wash conditions were 2X SSC, 1% SDS at 50°C . (Church and Gilbert (1984)). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequences of the cDNA inserts were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest sugar beet protox-1 cDNA obtained, designated "sugar beet protox-1", is full-length based on comparison with the other known plant protox peptide sequences (Table

1A). Sugar beet protox-1 is 1910-bp in length and encodes a protein of 60 kDa. The nucleotide sequence of this cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs:15 and 16, respectively. The sugar beet protein is 73% identical (82% similar) to the *Arabidopsis* protox-1 protein.

Sugar beet protox-1, in the pBluescript SK vector, was deposited July 29, 1996, as pWDC-16 (NRRL #B-21595N).

Example 5: Isolation of an Oilseed Rape Protox-1 cDNA Based on Sequence Homology to an *Arabidopsis* Protox-1 Coding Sequence

A Lambda Uni-Zap II cDNA library prepared from *Brassica napus* (3-4 wk. mature green leaves) was obtained from Dr. Guenther Ochs, Institut Fuer Allgemeine Botanik, Johannes Gutenberg-Universitaet Mainz, Germany (Günther Ochs, Gerald Schock, and Aloysius Wild, *Plant Physiol.* 103: 303-304 (1993)). Approximately 50,000 pfu of the cDNA library were plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The plaque lifts were probed with the *Arabidopsis* protox-1 cDNA (SEQ ID NO:1) labeled with ³²P-dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50° C. Wash conditions were 2X SSC, 1% SDS at 50° C. (Church and Gilbert (1984)). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequences of the cDNA inserts were determined by the chain termination method using dideoxy-terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest oilseed rape protox-1 cDNA obtained, designated "rape protox-1", is full-length based on comparison with the other known plant protox-peptide sequences (Table 1A). Rape protox-1 is 1784-bp in length and encodes a protein of 57.3kD. The nucleotide sequence of this cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs: 17 and 18, respectively. The oilseed rape protein is 87% identical (92% similar) to the *Arabidopsis* protox-1 protein.

Rape protox-1, in the pBluescript SK vector, was deposited August 23, 1996, as pWDC-17 (NRRL #B-21615).

Example 6: Isolation of a Rice Protox-1 cDNA Based on Sequence Homology to a Maize Protox-1 Coding Sequence

A Lambda gt11 cDNA library prepared from *Oryza sativa* (5 day etiolated shoots) was purchased from Clontech. Approximately 50,000 pfu of the cDNA library were plated at

a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5) labeled with ^{32}P -dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO_4 pH 7.0, 1 mM EDTA at 50°C . Wash conditions were 2X SSC, 1% SDS at 50°C . (Church and Gilbert (1984)). Positively hybridizing plaques were purified, and lambda DNA was prepared using the Wizard Lambda-Prep kit (Promega). The cDNA inserts were subcloned as *EcoRI* fragments into the pBluescript SK vector using standard techniques. The sequences of the cDNA inserts were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest rice protox-1 cDNA obtained, designated "rice protox-1", was 1224-bp in length. Rice protox-1 lacks coding sequence for the transit peptide plus approximately 172 amino acids of the mature coding sequence based on comparison with the other known plant protox peptide sequences (Table 1A). The nucleotide sequence of this partial cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs:19 and 20, respectively.

Rice protox-1, in the pBluescript SK vector, was deposited December 6, 1996, as pWDC-18 (NRRL #B-21648).

Example 7: Isolation of a Sorghum Protox-1 cDNA Based on Sequence Homology to a Maize Protox-1 Coding Sequence

A Lambda-Zap II cDNA library prepared from Sorghum bicolor (3-6 day green seedlings) was obtained from Dr. Klaus Pfizenmaier, Institute of Cell Biology and Immunology, University of Stuttgart, Germany (Harald Wajant, Karl-Wolfgang Mundry, and Klaus Pfizenmaier, *Plant Mol. Biol.* 26: 735-746 (1994)). Approximately 50,000 pfu of the cDNA library were plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5) labeled with ^{32}P -dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO_4 pH 7.0, 1 mM EDTA at 50°C . Wash conditions were 2X SSC, 1% SDS at 50°C . (Church and Gilbert (1984)). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequences of the cDNA inserts were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest sorghum protox-1 cDNA obtained, designated "sorghum protox-1", was 1590-bp in length. Sorghum protox-1 lacks coding sequence for the transit peptide plus approximately 44

amino acids of the mature coding sequence based on comparison with the other known plant protox peptide sequences (Table 1A). The nucleotide sequence of this partial cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs:21 and 22, respectively.

Sorghum protox-1, in the pBluescript SK vector, was deposited December 6, 1996, as pWDC-19 (NRRL #B-21649).

Example 8: Isolation of a Sugar Cane Protox-1 cDNA Based on Sequence Homology to a Maize Protox-1 Coding Sequence

A Lambda-Zap II cDNA library prepared from sugar cane was obtained from Henrik Albert of USDA/ARS at the Hawaii Agricultural Research Center. Approximately 50,000 pfu of the cDNA library were plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5) labeled with ^{32}P -dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO_4 pH 7.0, 1 mM EDTA at 50°C . Wash conditions were 2X SSC, 1% SDS at 50°C . (Church and Gilbert (1984)). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequences of the cDNA inserts were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest sugar cane protox-1 cDNA obtained, designated "sugar cane protox-1", was 633-bp in length. Sugar cane protox-1 lacks coding sequence for the transit peptide plus approximately 382 amino acids of the mature coding sequence based on comparison with the other known plant protox peptide sequences (Table 1A). The nucleotide sequence of this partial cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs:36 and 37, respectively.

Example 9: Demonstration of Plant Protox Clone Sensitivity to Protox Inhibitory Herbicides in a Bacterial System

Liquid cultures of protox-1/SASX38, protox-2/SASX38 and pBluescript/XL1-Blue were grown in L amp¹⁰⁰. One hundred microliter aliquots of each culture were plated on L amp¹⁰⁰ media containing various concentrations (1.0nM-10mM) of a protox inhibitory aryluracil herbicide of formula XVII. Duplicate sets of plates were incubated for 18 hours at 37°C .

The protox⁺ *E. coli* strain XL1-Blue showed no sensitivity to the herbicide at any concentration, consistent with reported resistance of the native bacterial enzyme to similar herbicides. The protox-1/SASX38 was clearly sensitive, with the lawn of bacteria almost entirely eliminated by inhibitor concentrations as low as 10nM. The protox-2/SASX38 was also sensitive, but only at a higher concentration (10µM) of the herbicide. The herbicide was effective even on plates maintained almost entirely in the dark. The toxicity of the herbicide was entirely eliminated by the addition of 20 µg/ml hematin to the plates.

The different herbicide tolerance between the two plant protox strains is likely the result of differential expression from these two plasmids, rather than any inherent difference in enzyme sensitivity. Protox-1/SASX38 grows much more slowly than protox-2/SASX38 in any heme-deficient media. In addition, the Mzprotox-2/SASX38 strain, with a growth rate comparable to Arab protox-1/SASX38, is also very sensitive to herbicide at the lower (10-100nM) concentrations.

Section B: Identification and Characterization of Plant Protox Genes Resistant to Protox-Inhibitory Herbicides

Example 10: Selecting for Plant Protox Genes Resistant to Protox-Inhibitory Herbicides in the *E. coli* Expression System

An *Arabidopsis thaliana* (Landsberg) cDNA library in the plasmid vector pFL61 (Minet *et al.*, *Plant J.* 2:417-422 (1992)) was obtained and amplified. The *E. coli* hemG mutant SASX38 (Sasarman *et al.*, *J. Gen. Microbiol.* 113:297(1979)) was obtained and maintained on L media containing 20µg/ml hematin (United States Biochemicals). The plasmid library was transformed into SASX38 by electroporation using the Bio-Rad Gene Pulser and the manufacturer's conditions. The electroporated cells were plated on L agar containing 100µg/ml ampicillin at a density of approximately 500,000 transformants/10cm plate. The cells were then incubated at 37°C for 40 hours in low light and selected for the ability to grow without the addition of exogenous heme. Heme prototrophs were recovered at a frequency of 400/10⁷ from the pFL61 library. Sequence analysis of twenty-two complementing clones showed that nine are of the type designated "protox-1," the protox gene expected to express a chloroplastic protox enzyme.

The pFL61 library is a yeast expression library, with the *Arabidopsis* cDNAs inserted bidirectionally. These cDNAs can also be expressed in bacteria. The protox cDNAs apparently initiate at an in-frame ATG in the yeast PGK 3' sequence approximately 10 amino acids 5' to the *NotI* cloning site in the vector and are expressed either from the *lacZ*

promoter 300bp further upstream or from an undefined cryptic bacterial promoter. Because protox-1 cDNAs that included significant portions of a chloroplast transit sequence inhibited the growth of the *E. coli* SASX38 strain, the clone with the least amount of chloroplast transit sequence attached was chosen for mutagenesis/herbicide selection experiments. This clone, pSLV19, contains only 17 amino acids of the putative chloroplast transit peptide, with the DNA sequence beginning at bp-151 of the *Arabidopsis* protox-1 cDNA (SEQ ID NO:1).

The plasmid pSLV19 was transformed into the random mutagenesis strain XL1-Red (Stratagene, La Jolla, CA). The transformation was plated on L media containing 50ug/ml ampicillin and incubated for 48 hours at 37°C. Lawns of transformed cells were scraped from the plates and plasmid DNA prepared using the Wizard Megaprep kit (Promega, Madison, WI). Plasmid DNA isolated from this mutator strain is predicted to contain approximately one random base change per 2000 nucleotides (see Greener *et al.*, *Strategies* 7(2):32-34 (1994).

The mutated plasmid DNA was transformed into the *hemG* mutant SASX38 (Sasarman *et al.*, *J. Gen. Microbiol.* 113:297 (1979) and plated on L media containing various concentrations of protox-inhibiting herbicide (formula XVII). The plates were incubated for 2 days at 37°C. Plasmid DNA was isolated from all colonies that grew in the presence of herbicide concentrations that effectively killed the wild type strain. The isolated DNA was then transformed into SASX38 and plated again on herbicide to ensure that the resistance observed was plasmid-borne. The protox coding sequence from plasmids passing this screen was excised by *NotI* digestion, recloned into an unmutagenized vector, and tested again for the ability to confer herbicide tolerance. The DNA sequence of protox cDNAs that conferred herbicide resistance was then determined and mutations identified by comparison with the wild type *Arabidopsis* protox-1 sequence (SEQ ID NO:1).

A single coding sequence mutant was recovered from the first mutagenesis experiment. This mutant leads to enhanced herbicide "resistance" only by increasing growth rate. It contains a C to A mutation at nucleotide 197 in SEQ ID NO:1 in the truncated chloroplast transit sequence of pSLV19, converting an ACG codon for threonine to an AAG codon for lysine at amino acid 56 of SEQ ID NO:2, and resulting in better complementation of the bacterial mutant. This plasmid also contains a silent coding sequence mutation at nucleotide 1059, with AGT (Ser) changing to AGC (Ser). This plasmid was designated pMut-1.

The pMut-1 plasmid was then transformed into the mutator XL1-Red strain as described above and the mutated DNA was isolated and plated on an herbicide

concentration that is lethal to the unmutagenized pMut-1 protox gene. Herbicide tolerant colonies were isolated after two days at 37°C and analyzed as described above. Multiple plasmids were shown to contain herbicide resistant protox coding sequences. Sequence analysis indicated that the resistant genes fell into two classes. One resistance mutation identified was a C to T change at nucleotide 689 in the *Arabidopsis* protox-1 sequence set forth in SEQ ID NO:1. This change converts a GCT codon for alanine at amino acid 220 of SEQ ID NO:2 to a GTT codon for valine, and was designated pAraC-1Val (see, Table 1B; sub-sequence 3).

A second class of herbicide resistant mutant contains an A to G change at nucleotide 1307 in the *Arabidopsis* protox-1 sequence. This change converts a TAC codon for tyrosine at amino acid 426 to a TGC codon for cysteine, and was designated pAraC-2Cys (see, Table 1B; sub-sequence 7).

A third resistant mutant has a G to A change at nucleotide 691 in the *Arabidopsis* protox-1 sequence. This mutation converts a GGT codon for glycine at amino acid 221 to an AGT codon for serine at the codon position adjacent to the mutation in pAraC-1. This plasmid was designated pAraC-3Ser (see, Table 1B; sub-sequence 4).

Resistant mutant pAraC-2Cys, in the pMut-1 plasmid, was deposited on November 14, 1994 under the designation pWDC-7 with the Agricultural Research Culture Collection and given the deposit designation NRRL #21339N.

Example 11: Additional Herbicide-Resistant Codon Substitutions at Positions Identified in the Random Screen

The amino acids identified as herbicide resistance sites in the random screen are replaced by other amino acids and tested for function and for herbicide tolerance in the bacterial system. Oligonucleotide-directed mutagenesis of the *Arabidopsis* protox-1 sequence is performed using the Transformer Site-Directed Mutagenesis Kit (Clontech, Palo Alto, CA). After amino acid changes are confirmed by sequence analysis, the mutated plasmids are transformed into SASX38 and plated on L-amp¹⁰⁰ media to test for function and on various concentrations of protox-inhibiting herbicide to test for tolerance.

This procedure is applied to the alanine codon at nucleotides 688-690 and to the tyrosine codon at nucleotides 1306-1308 of the *Arabidopsis* protox-1 sequence (SEQ ID NO:1). The results demonstrate that the alanine codon at nucleotides 688-690 can be changed to a codon for valine (pAraC-1Val), threonine (pAraC-1Thr), leucine (pAraC-1Leu), cysteine (pAraC-1Cys), or isoleucine (pAraC-1Ile) to yield an herbicide-resistant protox enzyme that retains function (see, Table 1B; sub-sequence 3). The results further

demonstrate that the tyrosine codon at nucleotides 1306-1308 can be changed to a codon for cysteine (pAraC-2Cys), isoleucine (pAraC-2Ile), leucine (pAraC-2Leu), threonine (pAraC-2Thr), methionine (pAraC-2Met), valine (pAraC-2Val), or alanine (pAraC-2Ala) to yield an herbicide-resistant protox enzyme that retains function (see, Table 1B; sub-sequence 7).

Example 12: Isolation of Additional Mutations that Increase Enzyme Function and/or Herbicide Tolerance of Previously Identified Resistant Mutants

Plasmids containing herbicide resistant protox genes are transformed into the mutator strain XL1-Red and mutated DNA is isolated as described above. The mutated plasmids are transformed into SASX38 and the transformants are screened on herbicide concentrations (formula XVII) sufficient to inhibit growth of the original "resistant" mutant. Tolerant colonies are isolated and the higher tolerance phenotype is verified as being coding sequence dependent as described above. The sequence of these mutants is determined and mutations identified by comparison to the progenitor sequence.

This procedure was applied to the pAraC-1Val mutant described above. The results demonstrate that the serine codon at amino acid 305 (SEQ ID NO:2) can be changed to a codon for leucine to yield an enzyme with higher tolerance to protox-inhibiting herbicides than the pAraC-1Val mutant alone. This second site mutation is designated AraC305Leu (see, Table 1B; sub-sequence 13). The same results are demonstrated for the threonine codon at amino acid 249, where a change to either isoleucine or to alanine leads to a more tolerant enzyme (see, Table 1B; sub-sequence 12). These changes are designated AraC249Ile and AraC249Ala, respectively.

The procedure was also applied to the pAraC-2Cys mutant described above. The results demonstrate that the proline codon at amino acid 118 (SEQ ID NO:2) can be changed to a codon for leucine to yield an enzyme with higher tolerance to protox-inhibiting herbicides than the pAraC-1Cys mutant alone. This mutation is designated AraC118Leu (see, Table 1B; sub-sequence 11). The same results are demonstrated for the serine codon at amino acid 305, where a change to leucine leads to a more tolerant pAraC-2Cys enzyme (see, Table 1B; sub-sequence 13). This change was also isolated with the pAraC-1Val mutant as described above and is designated AraC305Leu. Additional mutations that enhance the herbicide resistance of the pAraC-2Cys mutant include an asparagine to serine change at amino acid 425, designated AraC425Ser (Table 1B; sub-sequence 14), and a tyrosine to cysteine at amino acid 498, designated AraC498Cys (Table 1B; sub-sequence 15).

These changes (Table 1B; sub-sequences 11-15) are referred to as "second site" mutations, because they are not sufficient to confer herbicide tolerance alone, but rather enhance the function and/or the herbicide tolerance of an already mutant enzyme. This does not preclude the possibility that other amino acid substitutions at these sites could suffice to produce an herbicide tolerant enzyme since exhaustive testing of all possible replacements has not been performed.

Example 13: Combining Identified Resistance Mutations with Identified Second Site Mutations to Create Highly Functional/Highly Tolerant Protox Enzymes

The AraC305Leu mutation described above was found to enhance the function/herbicide resistance of both the AraC-1Val and the AraC-2Cys mutant plasmids. In an effort to test the general usefulness of this second site mutation, it was combined with the AraC-2Leu, AraC-2Val, and AraC-2Ile mutations and tested for herbicide tolerance. In each case, the AraC305Leu change significantly increased the growth rate of the resistant protox mutant on protox-inhibiting herbicide. Combinations of the AraC-2Ile resistant mutant with either the second site mutant AraC249Ile or AraC118Leu also produced more highly tolerant mutant protox enzymes. The AraC249Ile mutation demonstrates that a second site mutation identified as enhancing an AraC-1 (sub-sequence 3) mutant may also increase the resistance of an AraC-2 (sub-sequence 7) mutant. A three mutation plasmid containing AraC-2Ile, AraC305Leu, and AraC249Ile (Table 1B; sub-sequences 7, 13, and 12) has also been shown to produce a highly functional, highly herbicide tolerant protox-1 enzyme.

Example 14: Identification of Sites in the Maize Protox-1 Gene that Can Be Mutated to Give Herbicide Tolerance

The pMut-1 *Arabidopsis* protox -1 plasmid described above is very effective when used in mutagenesis/screening experiments in that it gives a high frequency of genuine coding sequence mutants, as opposed to the frequent up-promoter mutants that are isolated when other plasmids are used. In an effort to create an efficient plasmid screening system for maize protox-1, the maize cDNA was engineered into the pMut-1 vector in approximately the same sequence context as the *Arabidopsis* cDNA. Using standard methods of overlapping PCR fusion, the 5' end of the pMut-1 *Arabidopsis* clone (including 17 amino acids of chloroplast transit peptide with one mis-sense mutation as described above) was fused to the maize protox-1 cDNA sequence starting at amino acid number 14 of the maize sequence (SEQ ID NO:6). The 3' end of the maize cDNA was unchanged.

NotI restriction sites were placed on both ends of this fusion, and the chimeric gene was cloned into the pFL61 plasmid backbone from pMut-1. Sequence analysis revealed a single nucleotide PCR-derived silent mutation that converts the ACG codon at nucleotides 745-747 (SEQ ID NO:5) to an ACT codon, both of which encode threonine. This chimeric Arab-maize protox-1 plasmid was designated pMut-3.

The pMut-3 plasmid was transformed into the mutator XL1-Red strain as described above and the mutated DNA was isolated and plated on a herbicide concentration (formula XVII) that was lethal to the unmutagenized pMut-3 maize protox gene. Herbicide tolerant colonies were isolated after two days at 37°C and analyzed as described above. This analysis revealed multiple plasmids containing herbicide resistant protox coding sequences. Sequence analysis showed 5 single base changes that individually result in an herbicide tolerant maize protox-1 enzyme. Three of these mutations correspond to amino acid changes previously shown to confer tolerance at the homologous position in the *Arabidopsis* protox-1 gene. Two of the three are pMzC-1Val and pMzC-1Thr, converting the alanine (GCT) at amino acid 164 (SEQ ID NO:6) to either valine (GAT) or to threonine (ACT). This position corresponds to the pAraC-1 mutations described above (see, Table 1B; sub-sequence 3). The third analogous change, pMzC-3Ser, converts the glycine (GGT) at amino acid 165 to Serine (AGT), corresponding to the AraC-3Ser mutation described above (see, Table 1B; sub-sequence 4). These results serve to validate the expectation that herbicide-tolerant mutations identified in one plant protox gene may also confer herbicide tolerance in an equivalent plant protox gene from another species.

Two of the mutations isolated from the maize protox-1 screen result in amino acid changes at residues not previously identified as herbicide resistance sites. One change (Mz159Phe) converts cysteine (TGC) to phenylalanine (TTC) at amino acid 159 of the maize protox-1 sequence (SEQ ID NO:6) (see, Table 1B; sub-sequence 2). The second (Mz419Thr) converts isoleucine (ATA) to threonine (ACA) at amino acid 419 (see, Table 1B; sub-sequence 9).

Additional amino acid substitutions were made and tested at three of the maize mutant sites. Tolerance was demonstrated when glycine 165 was changed to leucine (pMzC-3Leu) or when cysteine 159 was changed to either leucine (Mz159Leu) or to lysine (Mz159Lys) (see, Table 1B; sub-sequences 4 and 2). Tolerant enzymes were also created by changing isoleucine 419 to histidine (Mz419His), glycine (Mz419Gly), or asparagine (Mz419Asn) (see, Table 1B; sub-sequence 9).

Individual amino acid changes that produced highly herbicide tolerant *Arabidopsis* protox-1 enzymes were engineered into the maize protox-1 gene by site-directed

mutagenesis as described above. Bacterial testing demonstrated that changing the alanine (GCT) at amino acid 164 (SEQ ID NO:6) to leucine (CTT) produced a highly tolerant maize enzyme (pMzC-1Leu) (see, Table 1B; sub-sequence 3). No mutation analogous to the AraC-2 site (Table 1B; sub-sequence 7) in *Arabidopsis* was isolated in the maize random screen. However, changing this site, tyrosine 370 in the maize enzyme (SEQ ID NO:6), to either isoleucine (pMzC-2Ile) or methionine (pMzC-2Met) did produce herbicide tolerant enzymes (see, Table 1B; sub-sequence 7).

Additional mutant screens performed as described earlier in this example, except using formulas XXIIIa and XXIIIb instead of XVII, identified three additional amino acid changes that confer tolerant protox enzymes. One, using formula XXIIIb, demonstrated that changing the arginine (CGT) at amino acid 88 (SEQ ID NO:6) to cysteine (TGT) produced a highly tolerant maize enzyme (Mz88Cys) (see, Table 1B; sub-sequence 1). Another, using formula IVc, demonstrated that changing the arginine (CGT) at amino acid 88 (SEQ ID NO:6) to leucine (CTT) produced a highly tolerant maize enzyme (Mz88Cys) (see, Table 1B; sub-sequence 1). Another, using formula XXIIIa, demonstrated that changing both the leucine (TTA) at amino acid 347 (SEQ ID NO:6) to serine (TCA) and the alanine (GCA) at amino acid 453 (SEQ ID NO:6) to threonine (ACA) produced a highly tolerant maize enzyme (Mz347Ser453Thr) (see, Table 1B; sub-sequences 16 and 17). Unlike the second site mutations described above, which increase enzyme function and/or herbicide tolerance of previously identified resistant mutants, Mz347Ser453Thr is a "double mutant" that requires that both mutations be present for herbicide tolerance.

Additional mutant screens performed as described earlier in this example identified two additional amino acid changes that confer tolerant protox enzymes. One, using formula XXIV, demonstrated that changing the alanine (GCA) at amino acid 175 (SEQ ID NO:6) to valine (GTA) or to threonine (ACA) produced a highly tolerant maize enzyme (Mz175Val and Mz175Thr, respectively) (see, Table 1B; sub-sequence 18). Another, using formula IVc, demonstrated that changing the leucine (TTG) at amino acid 337 (SEQ ID NO:6) to serine (TCG) produced a highly tolerant maize enzyme (Mz337Ser) (see, Table 1B; sub-sequence 19).

Example 15: Identification of Sites in the Wheat Protox-1 Gene that can be Mutated to Give Herbicide Tolerance

To create an efficient plasmid screening system for wheat protox-1, the wheat cDNA was engineered into the pMut-1 vector as described above for the maize cDNA. This chimeric Arab-wheat protox-1 plasmid is designated pMut-4. The pMut-4 DNA was

mutated and screened for herbicide tolerance as described above. This analysis revealed multiple plasmids containing herbicide resistant protox coding sequences. Sequence analysis showed 7 single base changes that individually result in an herbicide tolerant wheat protox-1 enzyme. Four of these mutations correspond to amino acid changes previously shown to confer tolerance at the homologous position in the *Arabidopsis* and/or in the maize protox-1 gene. Two, pWhtC-1Val and pWhtC-1Thr, convert the alanine (GCT) at amino acid 211 (SEQ ID NO:10) to valine (GAT) and to threonine (ACT), respectively. This position corresponds to the pAraC-1 mutations described above (see, Table 1B; sub-sequence 3). The third analogous change, pWhtC-3Ser, converts the glycine (GGT) at amino acid 212 to serine (AGT), corresponding to the AraC-3Ser mutation described above (see, Table 1B; sub-sequence 4). The fourth, Wht466Thr, converts isoleucine (ATA) to threonine (ACA) at amino acid 466, corresponding to the Mz419Thr mutant from maize (see, Table 1B; sub-sequence 9).

Three of the mutations isolated from the wheat protox-1 screen result in amino acid changes at residues not previously identified as herbicide resistance sites. One change (Wht356Leu) converts valine (GTT) to leucine (CTT) at amino acid 356 of the wheat protox-1 sequence (SEQ ID NO:10) (see, Table 1B; sub-sequence 6). A second (Wht421Pro) converts serine (TCT) to proline (CCT) at amino acid 421 (see, Table 1B; sub-sequence 8). The third (Wht502Ala) converts valine (GTT) to alanine (GCT) at amino acid 502 (see, Table 1B; sub-sequence 10).

Example 16: Identification of Sites in the Soybean Protox-1 Gene that can be Mutated to Give Herbicide Tolerance

To create an efficient plasmid screening system for soybean protox-1, the soybean cDNA was engineered into the pMut-1 vector as described above for the maize cDNA. This chimeric Arab-soybean protox-1 plasmid is designated pMut-5. The pMut-5 DNA was mutated and screened for herbicide tolerance as described above. This analysis revealed multiple plasmids containing herbicide resistant protox coding sequences. Sequence analysis showed 4 single base changes that individually result in an herbicide tolerant soybean protox-1 enzyme. Two of these mutations correspond to amino acid changes previously shown to confer tolerance at the homologous position in the *Arabidopsis* and/or in the wheat protox-1 gene. One, pSoyC-1Thr, converts the alanine (GCA) at amino acid 226 (SEQ ID NO:12) to threonine (ACA). This position corresponds to the pAraC-1Thr mutation described above (see, Table 1B; sub-sequence 3). The second analogous

change, Soy517Ala, converts the valine (GTT) at amino acid 517 to alanine (GCT), corresponding to the Wht502Ala mutation from wheat (see, Table 1B; sub-sequence 10).

Two of the mutations isolated from the soybean protox-1 screen result in amino acid changes at a residue not previously identified as an herbicide resistance site. One change (Soy369Ser) converts proline (CCT) to serine (TCT) at amino acid 369 of the soybean protox-1 sequence (SEQ ID NO:12) (see, Table 1B; sub-sequence 5). A second (Soy369His) converts this same proline369 to histidine (CAT) (see, Table 1B; sub-sequence 5).

Individual amino acid changes that produced highly herbicide tolerant *Arabidopsis* protox-1 enzymes were engineered into the soybean protox-1 gene by site directed mutagenesis as described above. Bacterial testing demonstrated that changing the alanine (GCA) at amino acid 226 (SEQ ID NO:12) to leucine (pSoyC-1Leu) produced a tolerant soybean enzyme (see, Table 1B; sub-sequence 3). Changing the tyrosine (TAC) at amino acid 432 (SEQ ID NO:12) to either leucine (pSoyC-2Leu) or isoleucine (pSoyC-2Ile) also produced herbicide tolerant enzymes (see, Table 1B; sub-sequence 7).

Example 17: Identification of Sites in the Sugar Beet Protox-1 Gene that can be Mutated to Give Herbicide Tolerance

To create an efficient plasmid screening system for sugar beet protox-1, the sugar beet cDNA was engineered into the pMut-1 vector as described above for the maize cDNA. This chimeric Arab-sugar beet protox-1 plasmid is designated pMut-6. The pMut-6 DNA was mutated and screened for herbicide tolerance as described above. This analysis revealed multiple plasmids containing herbicide resistant protox coding sequences. Sequence analysis showed a single base change that results in an herbicide tolerant sugar beet protox-1 enzyme. This change (pSugC-2Cys) converts tyrosine (TAC) at amino acid 449 to cysteine (TGC) and is analogous to the AraC-2 mutations in *Arabidopsis* (see, Table 1B; sub-sequence 7).

Individual amino acid changes that produced highly herbicide tolerant *Arabidopsis* protox-1 enzymes were engineered into the sugar beet protox-1 gene by site directed mutagenesis as described above. Bacterial testing demonstrated that changing the tyrosine (TAC) at amino acid 449 to leucine (pSugC-2Leu), isoleucine (pSugC-2Ile), valine (pSugC-2Val), or methionine (pSugC-2Met) produced herbicide tolerant sugar beet enzymes (see, Table 1B; sub-sequence 7).

Example 18: Identification of Sites in the Cotton Protox-1 Gene that can be Mutated to Give Herbicide Tolerance

In an effort to create an efficient plasmid screening system for cotton protox-1, the cotton cDNA was engineered into the pMut-1 vector as described above for the maize cDNA. This chimeric Arab-cotton protox-1 plasmid is designated pMut-7. The pMut-7 DNA was mutated and screened for herbicide tolerance as described above. This analysis revealed multiple plasmids containing herbicide resistant protox coding sequences. Sequence analysis showed 3 single base changes that individually result in an herbicide tolerant cotton protox-1 enzyme. Three mutants, pCotC-2Cys, pCotC-2His and pCotC-2Arg, change tyrosine (TAC) at amino acid 428 (SEQ ID NO:16) to cysteine (TGC), histidine (CAC) and to arginine (CGC), respectively (see, Table 1B; sub-sequence 7). Arginine is a novel substitution giving tolerance at this previously identified AraC-2 (sub-sequence 7) site. The third mutation (Cot365Ser) converts proline (CCT) to serine (TCT) at amino acid 365. This change corresponds to the soybean mutant Soy369Ser (see, Table 1B; sub-sequence 5).

Example 19: Demonstration of Resistant Mutations' Cross-Tolerance to Various Protox-Inhibiting Compounds

Resistant mutant plasmids, originally identified based on resistance against a single protox-inhibitory herbicide, were tested against a spectrum of other protox inhibiting compounds. For this test, the SASX38 strain containing the wild-type plasmid is plated on a range of concentrations of each compound to determine the lethal concentration for each one. Resistant mutant plasmids in SASX38 are plated and scored for the ability to survive on a concentration of each compound at least 10 fold higher than the concentration that is lethal to the SASX38 strain containing the wild-type plasmid.

Results from bacterial cross-tolerance testing, illustrated in Tables 3A and 3B, show that each of the mutations identified confers tolerance to a variety of protox inhibiting compounds.

Section C: Expression of Herbicide-Resistant Protox Genes in Transgenic Plants

Example 20: Engineering of Plants Tolerant to Protox-Inhibiting Herbicides by Homologous Recombination or Gene Conversion

Because the described mutant coding sequences effectively confer herbicide tolerance when expressed under the control of the native protox promoter, targeted

changes to the protox coding sequence in its native chromosomal location represent an alternative means for generating herbicide tolerant plants and plant cells. A fragment of protox DNA containing the desired mutations, but lacking its own expression signals (either promoter or 3' untranslated region) can be introduced by any of several art-recognized methods (for instance, *Agrobacterium* transformation, direct gene transfer to protoplasts, microprojectile bombardment), and herbicide-tolerant transformants selected. The introduced DNA fragment also contains a diagnostic restriction enzyme site or other sequence polymorphism that is introduced by site-directed mutagenesis in vitro without changing the encoded amino acid sequence (i.e. a silent mutation). As has been previously reported for various selectable marker and herbicide tolerance genes (see, e.g., Paszkowski *et al.*, *EMBO J.* 7: 4021-4026 (1988); Lee *et al.*, *Plant Cell* 2: 415-425 (1990); Risseuw *et al.*, *Plant J.* 7: 109-119 (1995)). some transformants are found to result from homologous integration of the mutant DNA into the protox chromosomal locus, or from conversion of the native protox chromosomal sequence to the introduced mutant sequence. These transformants are recognized by the combination of their herbicide-tolerant phenotype, and the presence of the diagnostic restriction enzyme site in their protox chromosomal locus.

Example 21: Construction of Plant Transformation Vectors

Numerous transformation vectors are available for plant transformation, and the genes of this invention can be used in conjunction with any such vectors. The selection of vector for use will depend upon the preferred transformation technique and the target species for transformation. For certain target species, different antibiotic or herbicide selection markers may be preferred. Selection markers used routinely in transformation include the *nptII* gene, which confers resistance to kanamycin and related antibiotics (Messing & Vierra, *Gene* 19: 259-268 (1982); Bevan *et al.*, *Nature* 304:184-187 (1983)), the *bar* gene, which confers resistance to the herbicide phosphinothricin (White *et al.*, *Nucl Acids Res* 18: 1062 (1990), Spencer *et al.* *Theor Appl Genet* 79: 625-631(1990)), the *hph* gene, which confers resistance to the antibiotic hygromycin (Blochinger & Diggelmann, *Mol Cell Biol* 4: 2929-2931), and the *dhfr* gene, which confers resistance to methotrexate (Bourouis *et al.*, *EMBO J.* 2(7): 1099-1104 (1983)).

I. Construction of Vectors Suitable for *Agrobacterium* Transformation

Many vectors are available for transformation using *Agrobacterium tumefaciens*. These typically carry at least one T-DNA border sequence and include vectors such as

pBIN19 (Bevan, *Nucl. Acids Res.* (1984)) and pXYZ. Below the construction of two typical vectors is described.

Construction of pCIB200 and pCIB2001: The binary vectors pCIB200 and pCIB2001 are used for the construction of recombinant vectors for use with *Agrobacterium* and was constructed in the following manner. pTJS75kan was created by *NarI* digestion of pTJS75 (Schmidhauser & Helinski, *J Bacteriol.* 164: 446-455 (1985)) allowing excision of the tetracycline-resistance gene, followed by insertion of an *AccI* fragment from pUC4K carrying an NPTII (Messing & Vierra, *Gene* 19: 259-268 (1982); Bevan *et al.*, *Nature* 304: 184-187 (1983); McBride *et al.*, *Plant Molecular Biology* 14: 266-276 (1990)). *XhoI* linkers were ligated to the *EcoRV* fragment of pCIB7, which contains the left and right T-DNA borders, a plant selectable *nos/nptII* chimeric gene and the pUC polylinker (Rothstein *et al.*, *Gene* 53: 153-161 (1987)), and the *XhoI*-digested fragment was cloned into *Sall*-digested pTJS75kan to create pCIB200 (see also EP 0 332 104, example 19). pCIB200 contains the following unique polylinker restriction sites: *EcoRI*, *SstI*, *KpnI*, *BglII*, *XbaI*, and *Sall*. pCIB2001 is a derivative of pCIB200, which is created by the insertion into the polylinker of additional restriction sites. Unique restriction sites in the polylinker of pCIB2001 are *EcoRI*, *SstI*, *KpnI*, *BglII*, *XbaI*, *Sall*, *MluI*, *BclI*, *AvrII*, *Apal*, *HpaI*, and *StuI*. pCIB2001, in addition to containing these unique restriction sites also has plant and bacterial kanamycin selection, left and right T-DNA borders for *Agrobacterium*-mediated transformation, the RK2-derived *trfA* function for mobilization between *E. coli* and other hosts, and the *OriT* and *OriV* functions also from RK2. The pCIB2001 polylinker is suitable for the cloning of plant expression cassettes containing their own regulatory signals.

Construction of pCIB10 and Hygromycin Selection Derivatives Thereof: The binary vector pCIB10 contains a gene encoding kanamycin resistance for selection in plants, T-DNA right and left border sequences and incorporates sequences from the wide host-range plasmid pRK252 allowing it to replicate in both *E. coli* and *Agrobacterium*. Its construction is described by Rothstein *et al.*, *Gene* 53: 153-161 (1987). Various derivatives of pCIB10 have been constructed that incorporate the gene for hygromycin B phosphotransferase described by Gritz *et al.*, *Gene* 25: 179-188 (1983)). These derivatives enable selection of transgenic plant cells on hygromycin only (pCIB743), or hygromycin and kanamycin (pCIB715, pCIB717).

II. Construction of Vectors Suitable for non-*Agrobacterium* Transformation.

Transformation without the use of *Agrobacterium tumefaciens* circumvents the requirement for T-DNA sequences in the chosen transformation vector and consequently

vectors lacking these sequences can be utilized in addition to vectors such as the ones described above that contain T-DNA sequences. Transformation techniques that do not rely on *Agrobacterium* include transformation via particle bombardment, protoplast uptake (e.g. PEG and electroporation) and microinjection. The choice of vector depends largely on the preferred selection for the species being transformed. Below, the construction of some typical vectors is described.

Construction of pCIB3064: pCIB3064 is a pUC-derived vector suitable for direct gene transfer techniques in combination with selection by the herbicide basta (or phosphinothricin). The plasmid pCIB246 comprises the CaMV 35S promoter in operational fusion to the *E. coli* GUS gene and the CaMV 35S transcriptional terminator and is described in the PCT published application WO 93/07278. The 35S promoter of this vector contains two ATG sequences 5' of the start site. These sites were mutated using standard PCR techniques in such a way as to remove the ATG's and generate the restriction sites *SspI* and *PvuII*. The new restriction sites were 96 and 37-bp away from the unique *Sall* site and 101 and 42-bp away from the actual start site. The resultant derivative of pCIB246 was designated pCIB3025. The GUS gene was then excised from pCIB3025 by digestion with *Sall* and *SacI*, the termini rendered blunt and religated to generate plasmid pCIB3060. The plasmid pJIT82 was obtained from the John Innes Centre, Norwich and the a 400-bp *SmaI* fragment containing the *bar* gene from *Streptomyces viridochromogenes* was excised and inserted into the *HpaI* site of pCIB3060 (Thompson *et al.* EMBO J 6: 2519-2523 (1987)). This generated pCIB3064, which comprises the *bar* gene under the control of the CaMV 35S promoter and terminator for herbicide selection, a gene for ampicillin resistance (for selection in *E. coli*) and a polylinker with the unique sites *SphI*, *PstI*, *HindIII*, and *BamHI*. This vector is suitable for the cloning of plant expression cassettes containing their own regulatory signals.

Construction of pSOG19 and pSOG35: pSOG35 is a transformation vector that utilizes the *E. coli* gene dihydrofolate reductase (DHFR) as a selectable marker conferring resistance to methotrexate. PCR was used to amplify the 35S promoter (~800-bp), intron 6 from the maize *Adh1* gene (~550-bp) and 18-bp of the GUS untranslated leader sequence from pSOG10. A 250-bp fragment encoding the *E. coli* dihydrofolate reductase type II gene was also amplified by PCR and these two PCR fragments were assembled with a *SacI*-*PstI* fragment from pBI221 (Clontech), which comprised the pUC19 vector backbone and the nopaline synthase terminator. Assembly of these fragments generated pSOG19, which contains the 35S promoter in fusion with the intron 6 sequence, the GUS leader, the DHFR gene and the nopaline synthase terminator. Replacement of the GUS leader in pSOG19

with the leader sequence from Maize Chlorotic Mottle Virus (MCMV) generated the vector pSOG35. pSOG19 and pSOG35 carry the pUC gene for ampicillin resistance and have *HindIII*, *SphI*, *PstI* and *EcoRI* sites available for the cloning of foreign sequences.

Example 22: Construction of Plant Expression Cassettes

Gene sequences intended for expression in transgenic plants are firstly assembled in expression cassettes behind a suitable promoter and upstream of a suitable transcription terminator. These expression cassettes can then be easily transferred to the plant transformation vectors described above in Example 21.

I. Promoter Selection

The selection of a promoter used in expression cassettes will determine the spatial and temporal expression pattern of the transgene in the transgenic plant. Selected promoters will express transgenes in specific cell types (such as leaf epidermal cells, mesophyll cells, root cortex cells) or in specific tissues or organs (roots, leaves or flowers, for example) and this selection will reflect the desired location of expression of the transgene. Alternatively, the selected promoter may drive expression of the gene under a light-induced or other temporally-regulated promoter. A further alternative is that the selected promoter be chemically regulated. This would provide the possibility of inducing expression of the transgene only when desired and caused by treatment with a chemical inducer.

II. Transcriptional Terminators

A variety of transcriptional terminators are available for use in expression cassettes. These are responsible for the termination of transcription beyond the transgene and its correct polyadenylation. Appropriate transcriptional terminators are those that are known to function in plants and include the CaMV 35S terminator, the *tm1* terminator, the nopaline synthase terminator, the pea *rbcS* E9 terminator, as well as terminators naturally associated with the plant protox gene (i.e. "protox terminators"). These can be used in both monocotyledons and dicotyledons.

III. Sequences for the Enhancement or Regulation of Expression

Numerous sequences have been found to enhance gene expression from within the transcriptional unit and these sequences can be used in conjunction with the genes of this invention to increase their expression in transgenic plants.

Various intron sequences have been shown to enhance expression, particularly in monocotyledonous cells. For example, the introns of the maize *Adh1* gene have been found to significantly enhance the expression of the wild-type gene under its cognate promoter when introduced into maize cells. Intron 1 was found to be particularly effective and enhanced expression in fusion constructs with the chloramphenicol acetyltransferase gene (Callis *et al.*, *Genes Develop.* 1: 1183-1200 (1987)). In the same experimental system, the intron from the maize *bronze1* gene had a similar effect in enhancing expression (Callis *et al.*, *supra*). Intron sequences have been routinely incorporated into plant transformation vectors, typically within the non-translated leader.

A number of non-translated leader sequences derived from viruses are also known to enhance expression, and these are particularly effective in dicotyledonous cells. Specifically, leader sequences from Tobacco Mosaic Virus (TMV, the "W-sequence"), Maize Chlorotic Mottle Virus (MCMV), and Alfalfa Mosaic Virus (AMV) have been shown to be effective in enhancing expression (*e.g.* Gallie *et al.* *Nucl. Acids Res.* 15: 8693-8711 (1987); Skuzeski *et al.* *Plant Molec. Biol.* 15: 65-79 (1990)).

IV. Targeting of the Gene Product Within the Cell

Various mechanisms for targeting gene products are known to exist in plants and the sequences controlling the functioning of these mechanisms have been characterized in some detail. For example, the targeting of gene products to the chloroplast is controlled by a signal sequence that is found at the amino terminal end of various proteins and that is cleaved during chloroplast import yielding the mature protein (*e.g.* Comai *et al.* *J. Biol. Chem.* 263: 15104-15109 (1988)). These signal sequences can be fused to heterologous gene products to effect the import of heterologous products into the chloroplast (van den Broeck *et al.* *Nature* 313: 358-363 (1985)). DNA encoding for appropriate signal sequences can be isolated from the 5' end of the cDNAs encoding the RUBISCO protein, the CAB protein, the EPSP synthase enzyme, the GS2 protein and many other proteins that are known to be chloroplast localized.

Other gene products are localized to other organelles such as the mitochondrion and the peroxisome (*e.g.* Unger *et al.* *Plant Molec. Biol.* 13: 411-418 (1989)). The cDNAs encoding these products can also be manipulated to effect the targeting of heterologous gene products to these organelles. Examples of such sequences are the nuclear-encoded ATPases and specific aspartate amino transferase isoforms for mitochondria. Targeting to cellular protein bodies has been described by Rogers *et al.*, *Proc. Natl. Acad. Sci. USA* 82: 6512-6516 (1985)).

In addition, sequences have been characterized that cause the targeting of gene products to other cell compartments. Amino terminal sequences are responsible for targeting to the ER, the apoplast, and extracellular secretion from aleurone cells (Koehler & Ho, *Plant Cell* 2: 769-783 (1990)). Additionally, amino terminal sequences in conjunction with carboxy terminal sequences are responsible for vacuolar targeting of gene products (Shinshi *et al.*, *Plant Molec. Biol.* 14: 357-368 (1990)).

By the fusion of the appropriate targeting sequences described above to transgene sequences of interest it is possible to direct the transgene product to any organelle or cell compartment. For chloroplast targeting, for example, the chloroplast signal sequence from the RUBISCO gene, the CAB gene, the EPSP synthase gene, or the GS2 gene is fused in frame to the amino terminal ATG of the transgene. The signal sequence selected should include the known cleavage site and the fusion constructed should take into account any amino acids after the cleavage site that are required for cleavage. In some cases this requirement may be fulfilled by the addition of a small number of amino acids between the cleavage site and the transgene ATG or alternatively replacement of some amino acids within the transgene sequence. Fusions constructed for chloroplast import can be tested for efficacy of chloroplast uptake by *in vitro* translation of *in vitro* transcribed constructions followed by *in vitro* chloroplast uptake using techniques described by (Bartlett *et al.* In: Edelman *et al.* (Eds.) *Methods in Chloroplast Molecular Biology*, Elsevier. pp. 1081-1091 (1982); Wasmann *et al.* *Mol. Gen. Genet.* 205: 446-453 (1986)). These construction techniques are well known in the art and are equally applicable to mitochondria and peroxisomes. The choice of targeting that may be required for expression of the transgenes will depend on the cellular localization of the precursor required as the starting point for a given pathway. This will usually be cytosolic or chloroplastic, although it may in some cases be mitochondrial or peroxisomal. The products of transgene expression will not normally require targeting to the ER, the apoplast or the vacuole.

The above described mechanisms for cellular targeting can be utilized not only in conjunction with their cognate promoters, but also in conjunction with heterologous promoters so as to effect a specific cell targeting goal under the transcriptional regulation of a promoter that has an expression pattern different to that of the promoter from which the targeting signal derives.

Example 23: Transformation of Dicotyledons

Transformation techniques for dicotyledons are well known in the art and include *Agrobacterium*-based techniques and techniques that do not require *Agrobacterium*. Non-

Agrobacterium techniques involve the uptake of exogenous genetic material directly by protoplasts or cells. This can be accomplished by PEG or electroporation mediated uptake, particle bombardment-mediated delivery, or microinjection. Examples of these techniques are described by Paszkowski *et al.*, *EMBO J* 3: 2717-2722 (1984), Potrykus *et al.*, *Mol. Gen. Genet.* 199: 169-177 (1985), Reich *et al.*, *Biotechnology* 4: 1001-1004 (1986), and Klein *et al.*, *Nature* 327: 70-73 (1987). In each case the transformed cells are regenerated to whole plants using standard techniques known in the art.

Agrobacterium-mediated transformation is a preferred technique for transformation of dicotyledons because of its high efficiency of transformation and its broad utility with many different species. The many crop species that are routinely transformable by *Agrobacterium* include tobacco, tomato, sunflower, cotton, oilseed rape, potato, soybean, alfalfa and poplar (EP 0 317 511 (cotton), EP 0 249 432 (tomato, to Calgene), WO 87/07299 (*Brassica*, to Calgene), US 4,795,855 (poplar)).

Transformation of the target plant species by recombinant *Agrobacterium* usually involves co-cultivation of the *Agrobacterium* with explants from the plant and follows protocols well known in the art. Transformed tissue is regenerated on selectable medium carrying the antibiotic or herbicide resistance marker present between the binary plasmid T-DNA borders.

Example 24: Transformation of Monocotyledons

Transformation of most monocotyledon species has now also become routine. Preferred techniques include direct gene transfer into protoplasts using PEG or electroporation techniques, and particle bombardment into callus tissue. Transformations can be undertaken with a single DNA species or multiple DNA species (*i.e.* co-transformation) and both these techniques are suitable for use with this invention. Co-transformation may have the advantage of avoiding complex vector construction and of generating transgenic plants with unlinked loci for the gene of interest and the selectable marker, enabling the removal of the selectable marker in subsequent generations, should this be regarded desirable. However, a disadvantage of the use of co-transformation is the less than 100% frequency with which separate DNA species are integrated into the genome (Schocher *et al.* *Biotechnology* 4: 1093-1096 (1986)).

Patent Applications EP 0 292 435 (to Ciba-Geigy), EP 0 392 225 (to Ciba-Geigy) and WO 93/07278 (to Ciba-Geigy) describe techniques for the preparation of callus and protoplasts from an elite inbred line of maize, transformation of protoplasts using PEG or electroporation, and the regeneration of maize plants from transformed protoplasts.

Gordon-Kamm *et al.*, *Plant Cell* 2: 603-618 (1990)) and Fromm *et al.*, *Biotechnology* 8: 833-839 (1990)) have published techniques for transformation of A188-derived maize line using particle bombardment. Furthermore, application WO 93/07278 (to Ciba-Geigy) and Koziel *et al.*, *Biotechnology* 11: 194-200 (1993)) describe techniques for the transformation of elite inbred lines of maize by particle bombardment. This technique utilizes immature maize embryos of 1.5-2.5 mm length excised from a maize ear 14-15 days after pollination and a PDS-1000He Biolistics device for bombardment.

Transformation of rice can also be undertaken by direct gene transfer techniques utilizing protoplasts or particle bombardment. Protoplast-mediated transformation has been described for *Japonica*-types and *Indica*-types (Zhang *et al.*, *Plant Cell Rep* 7: 379-384 (1988); Shimamoto *et al.* *Nature* 338: 274-277 (1989); Datta *et al.* *Biotechnology* 8: 736-740 (1990)). Both types are also routinely transformable using particle bombardment (Christou *et al.* *Biotechnology* 9: 957-962 (1991)).

Patent Application EP 0 332 581 (to Ciba-Geigy) describes techniques for the generation, transformation and regeneration of Pooideae protoplasts. These techniques allow the transformation of *Dactylis* and wheat. Furthermore, wheat transformation has been described by Vasil *et al.*, *Biotechnology* 10: 667-674 (1992)) using particle bombardment into cells of type C long-term regenerable callus, and also by Vasil *et al.*, *Biotechnology* 11: 1553-1558 (1993)) and Weeks *et al.*, *Plant Physiol.* 102: 1077-1084 (1993) using particle bombardment of immature embryos and immature embryo-derived callus. A preferred technique for wheat transformation, however, involves the transformation of wheat by particle bombardment of immature embryos and includes either a high sucrose or a high maltose step prior to gene delivery. Prior to bombardment, any number of embryos (0.75-1 mm in length) are plated onto MS medium with 3% sucrose (Murashige & Skoog, *Physiologia Plantarum* 15: 473-497 (1962)) and 3 mg/l 2,4-D for induction of somatic embryos, which is allowed to proceed in the dark. On the chosen day of bombardment, embryos are removed from the induction medium and placed onto the osmoticum (*i.e.* induction medium with sucrose or maltose added at the desired concentration, typically 15%). The embryos are allowed to plasmolyze for 2-3 h and are then bombarded. Twenty embryos per target plate is typical, although not critical. An appropriate gene-carrying plasmid (such as pCIB3064 or pSG35) is precipitated onto micrometer size gold particles using standard procedures. Each plate of embryos is shot with the DuPont Biolistics helium device using a burst pressure of ~1000 psi using a standard 80 mesh screen. After bombardment, the embryos are placed back into the dark to recover for about 24 h (still on osmoticum). After 24 hrs, the embryos are removed from

the osmoticum and placed back onto induction medium where they stay for about a month before regeneration. Approximately one month later the embryo explants with developing embryogenic callus are transferred to regeneration medium (MS + 1 mg/liter NAA, 5 mg/liter GA), further containing the appropriate selection agent (10 mg/l basta in the case of pCIB3064 and 2 mg/l methotrexate in the case of pSOG35). After approximately one month, developed shoots are transferred to larger sterile containers known as "GA7s" that contained half-strength MS, 2% sucrose, and the same concentration of selection agent. Patent application WO 94/13822 describes methods for wheat transformation and is hereby incorporated by reference.

Example 25: Isolation of the *Arabidopsis thaliana* Protox-1 Promoter Sequence

A Lambda Zap II genomic DNA library prepared from *Arabidopsis thaliana* (Columbia, whole plant) was purchased from Stratagene. Approximately 125,000 phage were plated at a density of 25,000 pfu per 15 cm Petri dish and duplicate lifts were made onto Colony/Plaque Screen membranes (NEN Dupont). The plaque lifts were probed with the *Arabidopsis* protox-1 cDNA (SEQ ID NO:1 labeled with 32P-dCTP by the random priming method (Life Technologies). Hybridization and wash conditions were at 65°C as described in Church and Gilbert, *Proc. Natl. Acad. Sci. USA* 81: 1991-1995 (1984). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. Sequence from the genomic DNA inserts was determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). One clone, AraPT1Pro, was determined to contain 580-bp of *Arabidopsis* sequence upstream from the initiating methionine (ATG) of the protox-1 protein coding sequence. This clone also contains coding sequence and introns that extend to-bp 1241 of the protox-1 cDNA sequence. The 580-bp 5' noncoding fragment is the putative *Arabidopsis* protox-1 promoter, and the sequence is set forth in SEQ ID NO:13.

AraPT1Pro was deposited December 15, 1995, as pWDC-11 (NRRL #B-21515)

Example 26: Construction of Plant Transformation Vectors Expressing Altered Protox-1 Genes Behind the Native *Arabidopsis* Protox-1 Promoter

A full-length cDNA of the appropriate altered *Arabidopsis* protox-1 cDNA was isolated as an *EcoRI-XhoI* partial digest fragment and cloned into the plant expression vector pCGN1761ENX (see Example 9 of International application no. PCT/IB95/00452 filed June 8, 1995, published Dec. 21, 1995 as WO 95/34659). This plasmid was digested with *NcoI* and *BamHI* to produce a fragment comprised of the complete protox-1 cDNA plus

a transcription terminator from the 3' untranslated sequence of the *tm1* gene of *Agrobacterium tumefaciens*. The AraPT1Pro plasmid described above was digested with *NcoI* and *BamHI* to produce a fragment comprised of pBluescript and the 580-bp putative *Arabidopsis* protox-1 promoter. Ligation of these two fragments produced a fusion of the altered protox cDNA to the native protox promoter. The expression cassette containing the protox-1 promoter/protox-1 cDNA/*tm1* terminator fusion was excised by digestion with *KpnI* and cloned into the binary vector pCIB200. The binary plasmid was transformed by electroporation into *Agrobacterium* and then into *Arabidopsis* using the vacuum infiltration method (Bechtold *et al.*, *C.R. Acad. Sci. Paris* 316: 1194-1199 (1993). Transformants expressing altered protox genes were selected on kanamycin or on various concentrations of protox inhibiting herbicide.

Example 27: Production of Herbicide Tolerant Plants by Expression of a Native Protox-1 Promoter/Altered Protox-1 Fusion

Using the procedure described above, an *Arabidopsis* protox-1 cDNA containing a TAC to ATG (Tyrosine to Methionine) change at nucleotides 1306-1308 in the protox-1 sequence (SEQ ID NO:1) was fused to the native protox-1 promoter fragment and transformed into *Arabidopsis thaliana*. This altered protox-1 enzyme (AraC-2Met) has been shown to be >10-fold more tolerant to various protox-inhibiting herbicides than the naturally occurring enzyme when tested in the previously described bacterial expression system. Seed from the vacuum infiltrated plants was collected and plated on a range (10.0nM-1.0uM) of a protox inhibitory aryluracil herbicide of formula XVII. Multiple experiments with wild-type *Arabidopsis* have shown that a 10.0nM concentration of this compound is sufficient to prevent normal seedling germination. Transgenic seeds expressing the AraC-2Met altered enzyme fused to the native protox-1 promoter produced normal *Arabidopsis* seedlings at herbicide concentrations up to 500nM, indicating at least 50-fold higher herbicide tolerance when compared to wild-type *Arabidopsis*. This promoter/altered protox enzyme fusion therefore functions as an effective selectable marker for plant transformation. Several of the plants that germinated on 100.0nM of protox-inhibiting herbicide were transplanted to soil, grown 2-3 weeks, and tested in a spray assay with various concentrations of the protox-inhibiting herbicide. When compared to empty vector control transformants, the AraPT1Pro/AraC-2Met transgenics were >10-fold more tolerant to the herbicide spray.

EXAMPLE 28: Demonstration of resistant mutations' cross-tolerance to various protox-inhibiting compounds in an *Arabidopsis* germination assay.

Using the procedure described above, an *Arabidopsis* protox-1 cDNA containing both a TAC to ATC (tyrosine to isoleucine) change at nucleotides 1306-1308 and a TCA to TTA (serine to leucine) change at nucleotides 945-947 in the protox-1 sequence (SEQ ID NO:1) was fused to the native protox-1 promoter fragment and transformed into *Arabidopsis thaliana*. This altered protox-1 enzyme (AraC-2Ile + AraC305Leu) has been shown to be >10-fold more tolerant to a protox inhibitory aryluracil herbicide of formula XVII than the naturally occurring enzyme when tested in a bacterial system (see Examples 9-13). Homozygous *Arabidopsis* lines containing this fusion were generated from transformants that showed high tolerance to a protox inhibiting herbicide in a seedling germination assay as described above. The seed from one line was tested for cross-tolerance to various protox-inhibitory compounds by repeating the germination assay on concentrations of the compounds that had been shown to inhibit germination of wild-type *Arabidopsis*. The results from these experiments are shown in Table 4.

Example 29: Isolation of a Maize Protox-1 Promoter Sequence

A Zea Mays (Missouri 17 inbred, etiolated seedlings) genomic DNA library in the Lambda FIX II vector was purchased from Stratagene. Approximately 250,000 pfu of the library was plated at a density of 50,000 phage per 15 cm plate and duplicate lifts were made onto Colony/Plaque screen membranes (NEN Dupont). The plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5) labeled with 32P-dCTP by the random priming method (Life Technologies). Hybridization and wash conditions were at 65°C as described in Church and Gilbert, *Proc. Natl. Acad. Sci. USA* 81: 1991-1995 (1984). Lambda phage DNA was isolated from three positively hybridizing phage using the Wizard Lambda Preps DNA Purification System (Promega). Analysis by restriction digest, hybridization patterns, and DNA sequence analysis identified a lambda clone containing approximately 3.5 kb of maize genomic DNA located 5' to the maize protox-1 coding sequence previously isolated as a cDNA clone. This fragment includes the maize protox-1 promoter. The sequence of this fragment is set forth in SEQ ID NO:14. From nucleotide 1 to 3532, this sequence is comprised of 5' noncoding sequence. From nucleotide 3533 to 3848, this sequence encodes the 5' end of the maize protox-1 protein.

A plasmid containing the sequence of SEQ ID NO:14 fused to the remainder of the maize protox-1 coding sequence was deposited March 19, 1996 as pWDC-14 (NRRL #B-21546).

Example 30: Construction of Plant Transformation Vectors Expressing Altered Protox-1 Genes Behind the Native Maize Protox-1 Promoter

The 3848-bp maize genomic fragment (SEQ ID NO:14) was excised from the isolated lambda phage clone as a *Sall*-*KpnI* partial digest product and ligated to a *KpnI*-*NotI* fragment derived from an altered maize protox-1 cDNA that contained an alanine to leucine change at amino acid 164 (SEQ ID NO:6). This created a fusion of the native maize protox-1 promoter to a full length cDNA that had been shown to confer herbicide tolerance in a bacterial system (Examples 9-14). This fusion was cloned into a pUC18 derived vector containing the CaMV 35S terminator sequence to create a protox promoter/altered protox cDNA/terminator cassette. The plasmid containing this cassette was designated pWCo-1.

A second construct for maize transformation was created by engineering the first intron found in the coding sequence from the maize genomic clone back into the maize cDNA. The insertion was made using standard overlapping PCR fusion techniques. The intron (SEQ ID NO:25) was 93-bp long and was inserted between nucleotides 203 and 204 of SEQ ID NO:6, exactly as it appeared in natural context in the lambda clone described in Example 29. This intron-containing version of the expression cassette was designated pWCo-2.

Example 31: Demonstration of Maize Protox-1 Promoter Activity in Transgenic Maize Plants

Maize plants transformed with maize protox promoter/altered protox fusions were identified using PCR analysis with primers specific for the transgene. Total RNA was prepared from the PCR-positive plants and reverse-transcribed using Superscript M-MLV (Life Technologies) under recommended conditions. Two microliters of the reverse transcription reaction was used in a PCR reaction designed to be specific for the altered protox sequence. While untransformed controls give no product in this reaction, approximately 85% of plants transformed with pWCo-1 gave a positive result, indicating the presence of mRNA derived from the transgene. This demonstrates some level of activity for the maize protox promoter. The RNA's from the transgenic maize plants were also subjected to standard northern blot analysis using the radiolabeled maize protox cDNA fragment from SEQ ID NO:6 as a probe. Protox-1 mRNA levels significantly above those of untransformed controls were detected in some of the transgenic maize plants. This elevated mRNA level is presumed to be due to expression of altered protox-1 mRNA from the cloned maize protox promoter.

Example 32: Isolation of a Sugar Beet Protox-1 Promoter Sequence

A genomic sugar beet library was prepared by Stratagene in the Lambda Fix II vector. Approximately 300,000 pfu of the library was plated and probed with the sugar beet protox-1 cDNA sequence (SEQ ID NO:17) as described for maize in Example 29. Analysis by restriction digest, hybridization patterns and DNA sequence analysis identified a lambda clone containing approximately 7 kb of sugar beet genomic DNA located 5' to the sugar beet coding sequence previously isolated as a cDNA clone. A *Pst*-*Sall* fragment of 2606-bp was subcloned from the lambda clone into a pBluescript vector. This fragment contains 2068-bp of 5' noncoding sequence and includes the sugar beet protox-1 promoter sequence. It also includes the first 453-bp of the protox-1 coding sequence and the 85-bp first intron contained in the coding sequence. The sequence of this fragment is set forth in SEQ ID NO:26.

A plasmid containing the sequence of SEQ ID NO:26 was deposited December 6, 1996 as pWDC-20 (NRRL #B-21650).

Example 33: Construction of Plant Transformation Vectors Expressing Altered Sugar Beet Protox-1 Genes Behind the Native Sugar Beet Protox-1 Promoter

The sugar beet genomic fragment (SEQ ID NO:26) was excised from the genomic subclone described in Example 32 as a *Sac*I-*Bsr*GI fragment that includes 2068-bp of 5' noncoding sequence and the first 300-bp of the sugar beet protox-1 coding sequence. This fragment was ligated to a *Bsr*GI-*Not*I fragment derived from an altered sugar beet protox-1 cDNA that contained a tyrosine to methionine change at amino acid 449 (SEQ ID NO:18). This created a fusion of the native sugar beet protox-1 promoter to a full length cDNA that had been shown to confer herbicide tolerance in a bacterial system (Examples 9-14). This fusion was cloned into a pUC18 derived vector containing the CaMV 35S terminator sequence to create a protox promoter/altered protox cDNA/terminator cassette. The plasmid containing this cassette was designated pWCo-3.

Example 34: Production of Herbicide Tolerant Plants by Expression of a Native Sugar Beet Protox-1 Promoter/Altered Sugar Beet Protox-1 Fusion

The expression cassette from pWCo-3 is transformed into sugar beet using any of the transformation methods applicable to dicot plants, including *Agrobacterium*, protoplast, and biolistic transformation techniques. Transgenic sugar beets expressing the altered

protox-1 enzyme are identified by RNA-PCR and tested for tolerance to protox-inhibiting herbicides at concentrations that are lethal to untransformed sugar beets.

Section D: Expression of Protox Genes in Plant Plastids

Example 35: Preparation of a Chimeric Gene Containing the Tobacco Plastid *clpP* Gene Promoter and Native *clpP* 5' Untranslated Sequence Fused to a GUS Reporter Gene and Plastid *rps16* Gene 3' Untranslated Sequence in a Plastid Transformation Vector

I. Amplification of the Tobacco Plastid *clpP* Gene Promoter and Complete 5' Untranslated RNA (5' UTR).

Total DNA from *N. tabacum* c.v. "Xanthi NC" was used as the template for PCR with a left-to-right "top strand" primer comprising an introduced *EcoRI* restriction site at position -197 relative to the ATG start codon of the constitutively expressed plastid *clpP* gene (primer Pclp_P1a: 5'-GCGGAATTCATACTTATTTATCATTAGAAAG-3' (SEQ ID NO:27); *EcoRI* restriction site underlined) and a right-to-left "bottom strand" primer homologous to the region from -21 to -1 relative to the ATG start codon of the *clpP* promoter that incorporates an introduced *NcoI* restriction site at the start of translation (primer Pclp_P2b: 5'-GCGGCATGGTAAATGAAAGAAAGAACTAAA-3' (SEQ ID NO:28); *NcoI* restriction site underlined). This PCR reaction was undertaken with *Pfu* thermostable DNA polymerase (Stratagene, La Jolla, CA) in a Perkin Elmer Thermal Cycler 480 according to the manufacturer's recommendations (Perkin Elmer/Roche, Branchburg, NJ) as follows: 7 min 95°C, followed by 4 cycles of 1 min 95°C / 2 min 43°C / 1 min 72°C, then 25 cycles of 1 min 95°C / 2 min 55°C / 1 min 72°C. The 213-bp amplification product comprising the promoter and 5' untranslated region of the *clpP* gene containing an *EcoRI* site at its left end and an *NcoI* site at its right end and corresponding to nucleotides 74700 to 74505 of the *N. tabacum* plastid DNA sequence (Shinozaki *et al.*, *EMBO J.* 5: 2043-2049 (1986)) was gel purified using standard procedures and digested with *EcoRI* and *NcoI* (all restriction enzymes were purchased from New England Biolabs, Beverly, MA).

II. Amplification of the Tobacco Plastid *rps16* Gene 3' Untranslated RNA Sequence (3'UTR).

Total DNA from *N. tabacum* c.v. "Xanthi NC" was used as the template for PCR as described above with a left-to-right "top strand" primer comprising an introduced *XbaI* restriction site immediately following the TAA stop codon of the plastid *rps16* gene encoding

ribosomal protein S16 (primer rps16P_1a (5'-
GCGTCTAGATCAACCGAAATTCAATTAAGG-3' (SEQ ID NO:30); *Xba*I restriction site
 underlined) and a right-to-left "bottom strand" primer homologous to the region from +134 to
 +151 relative to the TAA stop codon of *rps16* that incorporates an introduced *Hind*III
 restriction site at the 3' end of the *rps16* 3' UTR (primer rps16P_1b (5'-
CGCAAGCTTCAATGGAAGCAATGATAA-3' (SEQ ID NO:31); *Hind*III restriction site
 underlined). The 169-bp amplification product comprising the 3' untranslated region of the
rps16 gene containing an *Xba*I site at its left end and a *Hind*III site at its right end and
 containing the region corresponding to nucleotides 4943 to 5093 of the *N. tabacum* plastid
 DNA sequence (Shinozaki *et al.*, 1986) was gel purified and digested with *Xba*I and *Hind*III.

III. Ligation of a GUS Reporter Gene Fragment to the *clpP* Gene Promoter and 5' and 3' UTR's.

An 1864-bp β -glucuronidase (GUS) reporter gene fragment derived from plasmid pRAJ275 (Clontech) containing an *Nco*I restriction site at the ATG start codon and an *Xba*I site following the native 3' UTR was produced by digestion with *Nco*I and *Xba*I. This fragment was ligated in a four-way reaction to the 201-bp *Eco*RI/*Nco*I *clpP* promoter fragment, the 157-bp *Xba*I/*Hind*III *rps16* 3'UTR fragment, and a 3148-bp *Eco*RI/*Hind*III fragment from cloning vector pGEM3Zf(-) (Promega, Madison WI) to construct plasmid pPH138. Plastid transformation vector pPH140 was constructed by digesting plasmid pPRV111a (Zoubenko *et al.* 1994) with *Eco*RI and *Hind*III and ligating the resulting 7287-bp fragment to a 2222-bp *Eco*RI/*Hind*III fragment of pPH138.

Example 36: Preparation of a Chimeric Gene Containing the Tobacco Plastid *clpP* Gene Promoter Plus Tobacco Plastid *psbA* Gene Minimal 5' Untranslated Sequence Fused to a GUS Reporter Gene and Plastid *rps16* Gene 3' Untranslated Sequence in a Plastid Transformation Vector

Amplification of the tobacco plastid *clpP* gene promoter and truncated 5' untranslated RNA (5' UTR): Total DNA from *N. tabacum* c.v. "Xanthi NC" was used as the template for PCR as described above with the left-to-right "top strand" primer Pclp_P1a (SEQ ID NO:27) and a right-to-left "bottom strand" primer homologous to the region from -34 to -11 relative to the ATG start codon of the *clpP* promoter that incorporates an introduced *Xba*I restriction site in the *clpP* 5' UTR at position -11 (primer Pclp_P1b: 5'-
GCGTCTAGAAAGAACTAAATACTATATTTTCAC-3' (SEQ ID NO:29); *Xba*I restriction site underlined). The 202-bp amplification product comprising the promoter and truncated 5'

UTR of the *clpP* gene containing an *EcoRI* site at its left end and an *XbaI* site at its right end was gel purified and digested with *XbaI*. The *XbaI* site was subsequently filled in with Klenow DNA polymerase (New England Biolabs) and the fragment digested with *EcoRI*. This was ligated in a five-way reaction to a double stranded DNA fragment corresponding to the final 38 nucleotides and ATG start codon of the tobacco plastid *psbA* gene 5' UTR (with an *NcoI* restriction site overhang introduced into the ATG start codon) that was created by annealing the synthetic oligonucleotides minpsb_U (top strand: 5'-GGGAGTCCCTGATGATTAAATAAACCAAGATTTTAC-3' (SEQ ID NO:32)) and minpsb_L (bottom strand: 5'-CATGGTAAAATCTTGGTTTATTTAATCATCAGGGACTCCC-3' (SEQ ID NO:33); *NcoI* restriction site 5' overhang underlined), the *NcoI/XbaI* GUS reporter gene fragment described above, the *XbaI/HindIII* *rps16* 3'UTR fragment described above, and the *EcoRI/HindIII* pGEM3Zf(-) fragment described above to construct plasmid pPH139. Plastid transformation vector pPH144 was constructed by digesting plasmid pPRV111a (Zoubenko, *et al.*, *Nucleic Acids Res* 22: 3819-3824 (1994)) with *EcoRI* and *HindIII* and ligating the resulting 7287-bp fragment to a 2251-bp *EcoRI/HindIII* fragment of pPH139.

Example 37: Preparation of a Chimeric Gene Containing the Tobacco Plastid *clpP* Gene Promoter and Complete 5' Untranslated Sequence Fused to the *Arabidopsis thaliana* Protox-1 Coding Sequence and Plastid *rps16* Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

Miniprep DNA from plasmid AraC-2Met carrying an *Arabidopsis thaliana* *NotI* insert that includes cDNA sequences from the Protoporphyrinogen IX Oxidase ("protox") gene encoding a portion of the amino terminal plastid transit peptide, the full-length cDNA and a portion of the 3' untranslated region was used as the template for PCR as described above using a left-to-right "top strand" primer (with homology to nucleotides +172 to +194 relative to the ATG start codon of the full length precursor protein) comprising an introduced *NcoI* restriction site and new ATG start codon at the deduced start of the mature protox protein coding sequence (primer APRTXP1a: 5'-GGGACCATGGATTGTGTGATTGTCGGCGGAGG-3' (SEQ ID NO:34); *NcoI* restriction site underlined) and a right-to-left "bottom strand" primer homologous to nucleotides +917 to +940 relative to the native ATG start codon of the protox precursor protein (primer APRTXP1b: 5'-CTCCGCTCTCCAGCTTAGTGATAC-3' (SEQ ID NO:35)). The 778-bp product was digested with *NcoI* and *SfuI* and the resulting 682-bp fragment ligated to an 844-bp *SfuI/NotI* DNA fragment of AraC-2Met comprising the 3' portion of the protox coding sequence and a 2978-bp *NcoI/NotI* fragment of the cloning vector pGEM5Zf(+) (Promega,

Madison WI) to construct plasmid pPH141. Plastid transformation vector pPH143 containing the *clpP* promoter driving the Formula XVII-resistant AraC-2Met protox gene with the *rps16* 3' UTR was constructed by digesting pPH141 with *NcoI* and *SspI* and isolating the 1491-bp fragment containing the complete protox coding sequence, digesting the *rps16P_1a* and *rps16P_1b* PCR product described above with *HindIII*, and ligating these to a 7436-bp *NcoI/HindIII* fragment of pPH140.

Example 38: Preparation of a Chimeric Gene Containing the Tobacco Plastid *clpP* Gene Promoter Plus Tobacco Plastid *psbA* Gene Minimal 5' Untranslated Sequence Fused to the *Arabidopsis thaliana* Protox-1 Coding Sequence and Plastid *rps16* Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

Plastid transformation vector pPH145 containing the *clpP* promoter/*psbA* 5' UTR fusion driving the Formula XVII-resistant AraC-2Met protox gene with the *rps16* 3' UTR was constructed by digesting pPH141 with *NcoI* and *SspI* and isolating the 1491-bp fragment containing the complete protox coding sequence, digesting the *rps16P_1a* and *rps16P_1b* PCR product described above with *HindIII*, and ligating these to a 7465-bp *NcoI/HindIII* fragment of pPH144.

Example 39: Preparation of a Chimeric Gene Containing the Tobacco Plastid *clpP* Gene Promoter and 5' Untranslated Sequence Fused to the EPSP Synthase Coding Sequence and Plastid *rps16* Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

A cDNA library is screened for the 5-enolpyruvyl-3-phosphoshikimate synthase (EPSP synthase) gene (U.S. Patent Nos. 5,310,667, 5,312,910, and 5,633,435, all incorporated herein by reference). A plasmid clone containing the full length EPSP synthase gene cDNA is isolated by standard techniques of molecular cloning. PCR primers are designed for amplification of the mature-size EPSP synthase coding sequence from this plasmid using a top strand primer having a 5' extension containing an *NcoI* restriction site inserted at amino acid -1 from the deduced mature protein start, thus creating an ATG start codon at this position, and a bottom strand primer having a 5' extension containing an *XbaI* restriction site downstream of the stop codon of the EPSP mature coding sequence in the amplified PCR product. The PCR amplification is performed using the designated primers and plasmid DNA template according to standard protocols. Amplified products are cloned and sequenced and a *NcoI-XbaI* DNA fragment containing the complete mature EPSP

synthase coding sequence is isolated by restriction digest with *NcoI* and *XbaI*, electrophoresis on a 0.8% TAE agarose gel, and phenol extraction of the excised band.

A plastid transformation vector containing the *clpP* promoter directing transcription of the mature-sized EPSP synthase gene with the *rps16* 3' UTR is constructed by digesting pPH140 with *NcoI* and *XbaI* and purifying the fragment containing the vector backbone, 5' and 3' plastid integration targeting sequences, *aadA* selectable marker cassette, and *clpP* promoter / *rps16* 3' UTR expression sequences. This product is ligated in a two-way reaction with the *NcoI-XbaI* DNA fragment containing the mature-sized EPSP synthase coding sequence isolated as described above.

Example 40: Preparation of a Chimeric Gene Containing the Tobacco Plastid *clpP* Gene Promoter and 5' Untranslated Sequence Fused to the ALS Coding Sequence and Plastid *rps16* Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

A cDNA library is screened for the acetolactate synthase (ALS) gene (U.S. Patent No. 5,013,659). A plasmid clone containing the full length ALS gene cDNA is isolated by standard techniques of molecular cloning. PCR primers are designed for amplification of the mature-size ALS coding sequence from this plasmid using a top strand primer having a 5' extension containing an *NcoI* restriction site inserted at amino acid -1 from the deduced mature protein start, thus creating an ATG start codon at this position, and a bottom strand primer having a 5' extension containing an *XbaI* restriction site downstream of the stop codon of the ALS mature coding sequence in the amplified PCR product. The PCR amplification is performed using the designated primers and plasmid DNA template according to standard protocols. Amplified products are cloned and sequenced and a *NcoI-XbaI* DNA fragment containing the complete mature ALS coding sequence is isolated by restriction digest with *NcoI* and *XbaI*, electrophoresis on a 0.8% TAE agarose gel, and phenol extraction of the excised band.

A plastid transformation vector containing the *clpP* promoter driving the mature-sized ALS gene with the *rps16* 3' UTR is constructed by digesting pPH140 with *NcoI* and *XbaI* and purifying the fragment containing the vector backbone, 5' and 3' plastid integration targeting sequences, *aadA* selectable marker cassette, and *clpP* promoter / *rps16* 3' UTR expression sequences. This product is ligated in a two-way reaction with the *NcoI-XbaI* DNA fragment containing the mature-sized ALS coding sequence isolated as described above.

Example 41: Preparation of a Chimeric Gene Containing the Tobacco Plastid *clpP* Gene Promoter and 5' Untranslated Sequence Fused to the AHAS Coding Sequence and Plastid *rps16* Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

A cDNA library is screened for the acetohydroxyacid synthase (AHAS) gene (U.S. Patent No. 4,761,373). A plasmid clone containing the full length AHAS gene cDNA is isolated by standard techniques of molecular cloning. PCR primers are designed for amplification of the mature-size AHAS coding sequence from this plasmid using a top strand primer having a 5' extension containing an *NcoI* restriction site inserted at amino acid -1 from the deduced mature protein start, thus creating an ATG start codon at this position, and a bottom strand primer having a 5' extension containing an *XbaI* restriction site downstream of the stop codon of the AHAS mature coding sequence in the amplified PCR product. The PCR amplification is performed using the designated primers and plasmid DNA template according to standard protocols. Amplified products are cloned and sequenced and a *NcoI-XbaI* DNA fragment containing the complete mature AHAS coding sequence is isolated by restriction digest with *NcoI* and *XbaI*, electrophoresis on a 0.8% TAE agarose gel, and phenol extraction of the excised band.

A plastid transformation vector containing the *clpP* promoter driving the mature-sized AHAS gene with the *rps16* 3' UTR is constructed by digesting pPH140 with *NcoI* and *XbaI* and purifying the fragment containing the vector backbone, 5' and 3' plastid integration targeting sequences, *aadA* selectable marker cassette, and *clpP* promoter / *rps16* 3' UTR expression sequences. This product is ligated in a two-way reaction with the *NcoI-XbaI* DNA fragment containing the mature-sized AHAS coding sequence isolated as described above.

Example 42: Preparation of a Chimeric Gene Containing the Tobacco Plastid *clpP* Gene Promoter and 5' Untranslated Sequence Fused to the ACCase Coding Sequence and Plastid *rps16* Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

A cDNA library is screened for the acetylcoenzyme A carboxylase (ACCase) gene (U.S. Patent No. 5,162,602). A plasmid clone containing the full length ACCase gene cDNA is isolated by standard techniques of molecular cloning. PCR primers are designed for amplification of the mature-size ACCase coding sequence from this plasmid using a top strand primer having a 5' extension containing an *NcoI* restriction site inserted at amino acid -1 from the deduced mature protein start, thus creating an ATG start codon at this position, and a bottom strand primer having a 5' extension containing an *XbaI* restriction site

downstream of the stop codon of the ACCase mature coding sequence in the amplified PCR product. The PCR amplification is performed using the designated primers and plasmid DNA template according to standard protocols. Amplified products are cloned and sequenced and a *NcoI-XbaI* DNA fragment containing the complete mature ACCase coding sequence is isolated by restriction digest with *NcoI* and *XbaI*, electrophoresis on a 0.8% TAE agarose gel, and phenol extraction of the excised band.

A plastid transformation vector containing the *clpP* promoter driving the mature-sized ACCase gene with the *rps16* 3' UTR is constructed by digesting pPH140 with *NcoI* and *XbaI* and purifying the fragment containing the vector backbone, 5' and 3' plastid integration targeting sequences, *aadA* selectable marker cassette, and *clpP* promoter / *rps16* 3' UTR expression sequences. This product is ligated in a two-way reaction with the *NcoI-XbaI* DNA fragment containing the mature-sized ACCase coding sequence isolated as described above.

Example 43: Biolistic Transformation of the Tobacco Plastid Genome

Seeds of *Nicotiana tabacum* c.v. 'Xanthi nc' were germinated seven per plate in a 1" circular array on T agar medium and bombarded 12-14 days after sowing with 1 μ m tungsten particles (M10, Biorad, Hercules, CA) coated with DNA from plasmids pPH143 and pPH145 essentially as described in Svab, Z. and Maliga, P. (1993) *PNAS* 90, 913-917. Bombarded seedlings were incubated on T medium for two days after which leaves were excised and placed abaxial side up in bright light (350-500 μ mol photons/m²/s) on plates of RMOP medium (Svab, Z., Hajdukiewicz, P. and Maliga, P. (1990) *PNAS* 87, 8526-8530) containing 500 μ g/ml spectinomycin dihydrochloride (Sigma, St. Louis, MO). Resistant shoots appearing underneath the bleached leaves three to eight weeks after bombardment were subcloned onto the same selective medium, allowed to form callus, and secondary shoots isolated and subcloned. Complete segregation of transformed plastid genome copies (homoplasmy) in independent subclones was assessed by standard techniques of Southern blotting (Sambrook *et al.*, (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor). *BamHI/EcoRI*-digested total cellular DNA (Mettler, I. J. (1987) *Plant Mol Biol Reporter* 5, 346-349) was separated on 1% Tris-borate (TBE) agarose gels, transferred to nylon membranes (Amersham) and probed with ³²P-labeled random primed DNA sequences corresponding to a 0.7 kb *BamHI/HindIII* DNA fragment from pC8 containing a portion of the *rps7/12* plastid targeting sequence. Homoplasmy shoots are rooted aseptically on spectinomycin-containing MS/IBA medium (McBride, K. E. *et al.* (1994) *PNAS* 91, 7301-7305) and transferred to the greenhouse.

Example 44: Assessment of Herbicide Tolerance in Nt_pPH143 and Nt_pPH145 Plastid Transformant Lines

Primary homoplasmic transformant lines transformed with pPH143 (line Nt_pPH143) or with pPH145 (line Nt_pPH145), which were obtained as described in Example 43, were grown to maturity in the greenhouse. Flowers were either: (a) self-pollinated, (b) pollinated with wildtype tobacco (c.v. Xanthi nc), or (c) used as pollen donors to fertilize emasculated flowers of wildtype Xanthi plants. Plastid segregation of the linked spectinomycin resistance marker was verified by uniparental female inheritance of the spectinomycin-resistance phenotype in each transformant line using a minimum of 50 seeds per selection pool derived from either selfed or backcross capsules. Additional self or wildtype backcross (Xanthi pollen parent) seeds were germinated in soil. 36 plants of each line (143 1B-1, 143 1B-4, 143 4A-2, 143 4A-5, 145 7A-5, 145 7A-6, 145 8A-3) plus 36 wildtype Xanthi plants as isogenic controls were grown in separate 6" clay pots in a controlled environment cubicle. In order to assess tolerance to the protox inhibitor Formula XVII, plants of Xanthi and the seven transformant lines were distributed into eight identical 16-pot flats (2 plants of each type per flat). The flats were sprayed with Formula XVII until runoff at concentrations of either 0, 0.5, 2.5, 5, 10, 25, 50, or 100 mg Formula XVII per liter. Solutions were made up in water using 4 g/liter or 40 g/liter stock solutions of Formula XVII dissolved in dimethylsulfoxide (DMSO) and used immediately after preparation. Twenty microliters of the wetting agent Silwett was added to each 200 ml volume of herbicide solution for a final concentration of 0.01%. Flats were sprayed in the late afternoon and allowed to dry overnight before transfer to the growth cubicle. Tolerance was assessed by comparing leaf damage and wilting to the untransformed Xanthi controls at 0, 18 hrs, 48 hrs, and 6 days post-application. Severe damage was apparent on the Xanthi plants at all concentrations above 0.5 mg/l, and complete wilting/burn down occurred above 2.5 mg/l. Only slight damage occurred on the Nt_pPH143 plants even at the highest concentration (100 mg/liter), and the plants soon outgrew the bleached spots (the appearance of Xanthi at 0.5 mg/liter was approximately equivalent to Nt_pPH143 1B-1 at 100 mg/liter, giving a tolerance of ca. 200-fold).

Example 45: Plastid Transformation of Maize

Type I embryogenic callus cultures (Green *et al.* (1983) in A. Fazelahmad, K. Downey, J. Schultz, R.W. Voellmy, eds. *Advances in Gene Technology: Molecular Genetics of Plants and Animals*. Miami Winter Symposium Series, Vol. 20. Academic Press,

N.Y.) of the proprietary genotypes CG00526 and CG00714 are initiated from immature embryos, 1.5 - 2.5 mm in length, from greenhouse grown material. Embryos are aseptically excised from surface-sterilized ears approximately 14 days after pollination. Embryos of CG00526 are placed on D callus initiation media with 2% sucrose and 5mg/L chloramben (Duncan *et al.* (1985) *Planta* 165: 322-332) while those of CG00714 are placed onto KM callus initiation media with 3% sucrose and 0.75mg/L 2,4-d (Kao and Michayluk (1975) *Planta* 126, 105-110). Embryos and embryogenic cultures are subsequently cultured in the dark. Embryogenic responses are removed from the explants after ~14 days. CG00526 responses are placed onto D callus maintenance media with 2% sucrose and 0.5mg/L 2,4-d while those of CG00714 are placed onto KM callus maintenance media with 2% sucrose and 5mg/L Dicamba. After 3 to 8 weeks of weekly selective subculture to fresh maintenance media, high quality compact embryogenic cultures are established. Actively growing embryogenic callus pieces are selected as target tissue for gene delivery. The callus pieces are plated onto target plates containing maintenance medium with 12% sucrose approximately 4 hours prior to gene delivery. The callus pieces are arranged in circles, with radii of 8 and 10mm from the center of the target plate. Plasmid DNA is precipitated onto gold microcarriers as described in the DuPont Biolistics manual. Two to three μ g of each plasmid is used in each 6 shot microcarrier preparation. Genes are delivered to the target tissue cells using the PDS-1000He Biolistics device. The settings on the Biolistics device are as follows: 8 mm between the rupture disc and the macrocarrier, 10 mm between the macrocarrier and the stopping screen and 7 cm between the stopping screen and the target. Each target plate is shot twice using 650psi rupture discs. A 200 X 200 stainless steel mesh (McMaster-Carr, New Brunswick, NJ) is placed between the stopping screen and the target tissue.

Five days later, the bombed callus pieces are transferred to maintenance medium with 2% sucrose and 0.5mg/L 2,4-d, but without amino acids, and containing 750 or 1000 nM Formula XVII. The callus pieces are placed for 1 hour on the light shelf 4-5 hours after transfer or on the next day, and stored in the dark at 27°C for 5-6 weeks. Following the 5-6 week primary selection stage, yellow to white tissue is transferred to fresh plates containing the same medium supplemented with 500 or 750 nM Formula XVII. 4-5 hours after transfer or on the next day, the tissues are placed for 1 hour on the light shelf and stored in the dark at 27°C for 3-4 weeks. Following the 3-4 week secondary selection stage, the tissues are transferred to plates containing the same medium supplemented with 500 nM Formula XVII. Healthy growing tissue is placed for 1 hour on the light shelf and stored in the dark at 27°C. It is subcultured every two weeks until the colonies are large enough for regeneration.

At that point, colonies are transferred to a modified MS medium (Murashige and Skoog (1962) *Physiol. Plant* 15: 473-497) containing 3% sucrose (MS3S) with no selection agent and placed in the light. For CG00526, 0.25mg/L ancymidol and 0.5mg/L kinetin are added to this medium to induce embryo germination, while for CG00714, 2mg/L benzyl adenine is added. Regenerating colonies are transferred to MS3S media without ancymidol and kinetin, or benzyl adenine, for CG00526 or CG00714, respectively, after 2 weeks. Regenerating shoots with or without roots are transferred to boxes containing MS3S medium and small plants with roots are eventually recovered and transferred to soil in the greenhouse.

Example 46: Determination of the NH₂-terminus of the mature maize
protoporphyrinogen oxidase

Maize seedlings are grown 6 days at 27°C in the dark and illuminated 30 min prior to recovering etioplasts. Leaves are homogenised in extraction buffer (20 mM TES, 10 mM Hepes, 0.5 M sucrose, 1 mM MgCl₂, 1 mM EDTA, 0.2% BSA, pH 7.7, ±5 mM DTT) and filtered through cheesecloth layers. After centrifugation (7 min x 10000g) the pellet is resuspended in the same buffer and centrifuged 4 min x 150g. The supernatant is centrifuged 10 min x 6000g, and the pellet resuspended in a small amount of buffer.

For immunoprecipitation, the chloroplast fraction is diluted 5 fold in 20mM K-phosphate buffer, pH 7.0, 0.2 % Triton X-100, put on ice for 30 min and sonicated (3 x 6 sec) before being mixed with Protein A Sepharose CL-4B (Pharmacia). The supernatant is recovered by centrifugation and incubated with Protein A Sepharose CL-4B coupled to purified anti- protoporphyrinogen oxidase antibodies (anti- protox-1). The resin is washed with 20 mM Na-Phosphate buffer, pH 7.0, and specifically bound proteins are eluted by boiling in 3% SDS. Detergent is precipitated at 4°C, the supernatant lyophilised by speed-vacuum and resuspended in 20 mM Na-Phosphate buffer, pH 7.0.

Proteins are separated by SDS-electrophoresis according to Laemmli (1970) *Nature (London)* 227, 680-685, and silver stained using a Biorad (Hercules, CA) kit. The proteins separated by SDS-electrophoresis are blotted onto nitrocellulose (Millipore). Purified anti-protoporphyrinogen oxidase antibodies are diluted 1:15,000 before use. Anti rabbit IgG alkaline phosphatase conjugates (Sigma) are used as secondary antibodies and protoporphyrinogen oxidase bands are stained by using a nitroblue tetrazolium/ 5-bromo-4-chloro-3-indolyl phosphate kit (Biorad).

Protein samples for NH₂-sequencing are prepared as previously described (de Marco, A., Guzzardi, P., & Jamet, E. (1999) *Plant Physiol.* 120, 371-381), and Edman

analysis is carried out in a model 477A protein sequencer (Applied Biosystems Inc., Foster City, CA) according to the recommendations of the manufacturer. The NH₂-terminus of the mature maize protoporphyrinogen oxidase (protox-1) is found to be located at the alanine at position 1 in the amino acid sequence set forth in SEQ ID NO:45.

Purified recombinant protoporphyrinogen oxidase (see e.g. this application) is used to raise antibodies in rabbit using methods well known in the art. The antibodies are purified from sera using EMDTA Fractogel (Merck).

Example 47: Recombinant expression of the mature maize protoporphyrinogen oxidase

A. Construction of expression vectors and cell transformation.

The pET16b and pET28 bacterial expression vector (Novagen, Madison, USA), based on the T7 promoter driven system and containing a 5' His-tag is used. The coding sequence of the mature maize protox-1 is amplified by PCR. The PCR product is cloned into the BamHI site of pET16b. A plasmid containing the correct orientation of the insert is identified by sequencing and used for transformation of competent BL21 (DE3) and BL21 (DL3)pLysS cells.

B. Recombinant protein isolation.

E. coli cells, transformed with the vector hosting the mature maize protoporphyrinogen oxidase coding sequence, are cultured 16 h at 36°C under constant agitation. Isopropyl-β-D-thiogalactopyranoside (IPTG) is added to the medium 2 h before bacterial recovery. Alternative growth conditions are also used to increase the accumulation of active protein. First, induction of heat-shock proteins is stimulated by culturing the bacteria initially at 30°C and then increasing to 42°C in the presence of 3% ethanol. Second, the uptake of the compatible solute glycyl betaine is performed using sorbitol as an osmotic. Bacterial cells are recovered by centrifugation (15 min x 7500g) and stored at -80°C. The bacterial cells are resuspended in extraction buffer (40 mM Na-Phosphate, pH 7.2, 300 mM NaCl, 20% glycerol, 0.2% Triton X-100, 20 mM imidazole), sonicated (6 x 6 sec), and then centrifuged (25 min x 100000g). The supernatant is mixed with 2 mL of pre-washed immobilized nickel affinity chromatography resin (Talon, Clontech Lab.) in a 50 mL Falcon tube and gently rocked for 10 min, at 4°C. After centrifugation (2 min x 700 rpm), the supernatant is removed and the resin washed in extraction buffer (2X). Different protocols are used to complete the protein purification and described below.

Method 1): The resin is used to load a column and subsequently washed with 10 volumes of washing buffer (50 mM EPPS, pH 7.2, 20 % glycerol, 0.1% TritonX -100). Bound proteins are eluted with the washing buffer containing 100 mM EDTA, desalted and buffer

exchanged (50 mM EPPS, pH 8.5, 1 mM EDTA) using PD-10 Sephadex G-25 columns (Pharmacia), and loaded onto an ion-exchange Mono-Q FPLC column (Pharmacia). Proteins are eluted using a 0-1 M NaCl gradient. Fractions corresponding to the different proteins are collected separately, concentrated by speed-vacuum and dialysed against washing buffer. Alternatively, fractions recovered by affinity chromatography are buffer exchanged with 20 mM EPPS, pH 7.2, 150 mM NaCl and are loaded (~ 150 μ L corresponding to 150 ng protein) onto a pre-equilibrated gel-filtration Superdex 75 column (Pharmacia).

Method 2): Proteins are eluted as in Method 1), buffer exchanged against washing buffer, loaded onto a column, washed twice in ATP buffer (50 mM EPPS, pH 7.2, 20 % glycerol, 5 mM ATP, 7 mM MgCl₂, 0.1% TritonX -100) and eluted with imidazole buffer (50 mM EPPS, pH 7.2, 20% glycerol, 0.1% TritonX -100, 1 M imidazole).

Method 3): After the initial extraction buffer washing steps, the metal affinity resin is resuspended in ATP buffer and agitated for 30 min at 4 °C. The resin is recovered by centrifugation and washed again in the ATP buffer before being loaded onto a column and re-equilibrated with 50 mM EPPS, pH 7.2, 20% glycerol, 0.1% TritonX -100. Protein is eluted with imidazole buffer (described above). In some extractions, phospholipids (Sigma, St. Louis) were added to the elution buffer (L- α -phosphatidylcholine; L- α -phosphatidylcholine, dioleoyl; L- α -phosphatidyl-ethanolamine; L- α -phosphatidyl-L-serine; L- α -phosphatidylinositol; L- α -lysophosphatidylinositol; L- α -lysophosphatidylcholine; L- α -lysophosphatidyl-L-serine; L- α -lysophosphatidylethanolamine).

Resuspension of heat denaturated protein and of protein from homogenate pellet is performed using extraction buffer or re-equilibration buffer containing either phospholipides or chaperones. Proteins not immediately used for enzymatic tests are stored at -80 °C. The His-tag is removed by incubation in the presence of Factor Xa (New England Biolabs), according to manufacturer's instructions.

Table 1A

Alignment of the full-length and partial protox-1 amino acid sequences from *Arabidopsis* ("Arabpt-1"; SEQ ID NO:2), maize ("Mzpt-1"; SEQ ID NO:6), wheat ("Wtpt-1"; SEQ ID NO:10), soybean ("Soybeanpt-1"; SEQ ID NO:12), cotton ("Cottonpt-1"; SEQ ID NO:16), sugar beet ("Sugpt-1"; SEQ ID NO:18), oilseed rape ("Rapept-1"; SEQ ID NO:20), rice ("Ricept-1"; SEQ ID NO:22), sorghum ("Sorghumpt-1"; SEQ ID NO:24), and sugar cane ("Scpt-1"; SEQ ID NO:37). Alignment was performed using the PileUp program (GCG

package, University of Wisconsin, Madison, WI). Positions that may be modified according to the teachings herein to confer or enhance inhibitor resistance are shown in bold type.

	1					50
Rapept-1	MDLSLLRP..	QPFLSPFSNP	FPRSRPYKPL	
Arabpt-1	MELSLLRPTT	QSLFPSFSKP	NLRLNVYKPL	
Sorghumpt-1	
Mzpt-1	
Wtpt-1M	ATATVAAASP	LRGRVTGRPH	
Ricept-1	
Cottonpt-1MTAL	IDLSLLRSSP	SVSPFSIPHH	QHPPRFRKPF	
Soybeanpt1MV	SVFNEILFPP	NQTLLRPSLH	SPTSFFTSPT	RKFPRSRPNP	
Sugpt-1	MKSMALSNCI	PQTQCMPLRS	SGHYRGNCIM	LSIPCSLIGR	RGYYSHKKRR	
Scpt-1	
	51					100
Rapept-1	NLRCSVSGGS	VVGSSTIEGG	GGGKTVTADC	VIVGGGISGL	CIAQALVTKH	
Arabpt-1	RLRCSVAGGP	TVGSSKIEGG	GGT.TITDC	VIVGGGISGL	CIAQALATKH	
Sorghumpt-1	
Mzpt-1ADC	VVGGGISGL	CTAQALATRH	
Wtpt-1	RVRPRCATAS	SATETPAAPG	VRL...SAEC	VIVGAGISGL	CTAQALATRY	
Ricept-1	
Cottonpt-1	KLRCSLAEGP	TISSSKIDGG	ESS...IADC	VIVGGGISGL	CIAQALATKH	
Soybeanpt1	ILRCSIAEES	TASPPKTR..	DSA...PVDC	VVGGGVSGSL	CIAQALATKH	
Sugpt-1	MSMSCSTSSG	SKSAVKEAGS	GSGAGLLDC	VIVGGGISGL	CIAQALCTKH	
Scpt-1	

	101				150
Rapept-1	PDA..AKNVM	VTEAKDRVGG	NIIT..REEQ	GFLWEEGPNS	FQPSDPMLTM
Arabpt-1	PDA..APNLI	VTEAKDRVGG	NIIT..REEN	GFLWEEGPNS	FQPSDPMLTM
Sorghumpt-1STVERPEE	GYLWEEGPNS	FQPSDPVLSM
Mzpt-1	..G..VGDVL	VTEARARPGG	NITTVERPEE	GYLWEEGPNS	FQPSDPVLTM
Wtpt-1	..G..VSDLL	VTEARDRPGG	NITTVERPDE	GYLWEEGPNS	FQPSDPVLTM
Ricept-1
Cottonpt-1	RDV..ASNVI	VTEARDRVGG	NITTVER..D	GYLWEEGPNS	FQPSDPILTM
Soybeanpt1	..A..NANVV	VTEARDRVGG	NITTMER..D	GYLWEEGPNS	FQPSDPMLTM
Sugpt-1	SSSSLSPNFI	VTEAKDRVGG	NIVTVE..AD	GYIWEEGPNS	FQPSDAVLTM
Scpt-1
	151				200
Rapept-1	VVDSGLKDDL	VLGDPTAPRF	VLWNGKLRPV	PSKLTDLFFF	DLMSIGGKIR
Arabpt-1	VVDSGLKDDL	VLGDPTAPRF	VLWNGKLRPV	PSKLTDLFFF	DLMSIGGKIR
Sorghumpt-1	AVDSGLKDDL	VFGDPNAPRF	VLWEGKLRPV	PSKPADLFFF	DLMSIPGKLR
Mzpt-1	AVDSGLKDDL	VFGDPNAPRF	VLWEGKLRPV	PSKPADLFFF	DLMSIPGKLR
Wtpt-1	AVDSGLKDDL	VFGDPNAPRF	VLWEGKLRPV	PSKPGDLFFF	SLMSIPGKLR
Ricept-1
Cottonpt-1	AVDSGLKDDL	VLGDPNAPRF	VLWEGKLRPV	PSKPTDLFFF	DLMSIAGKLR
Soybeanpt1	VVDSGLKDEL	VLGDPDAPRF	VLWNRKLRPV	PGKLTDLFFF	DLMSIGGKIR
Sugpt-1	AVDSGLKDEL	VLGDPNAPRF	VLWNDKLRPV	PSSLTDLFFF	DLMTIPGKIR
Scpt-1
	201				250
Rapept-1	AGFGAIGIRP	SPPGREESVE	EFVRRNLGDE	VFERLIEPFC	SGVYAGDPAK
Arabpt-1	AGFGALGIRP	SPPGREESVE	EFVRRNLGDE	VFERLIEPFC	SGVYAGDPSK
Sorghumpt-1	AGLGALGIRP	PAPGREESVE	EFVRRNLGAE	VFERLIEPFC	SGVYAGDPSK
Mzpt-1	AGLGALGIRP	PPPGREESVE	EFVRRNLGAE	VFERLIEPFC	SGVYAGDPSK
Wtpt-1	AGLGALGIRP	PPPGREESVE	EFVRRNLGAE	VFERLIEPFC	SGVYAGDPSK
Ricept-1
Cottonpt-1	AGFGAIGIRP	PPPGYEEESVE	EFVRRNLGAE	VFERFIEPFC	SGVYAGDPSK
Soybeanpt1	AGFGALGIRP	PPPGHEESVE	EFVRRNLGDE	VFERLIEPFC	SGVYAGDPSK
Sugpt-1	AALGALGFRP	SPPPHEESVE	HFVRRNLGDE	VFERLIEPFC	SGVYAGDPAK
Scpt-1
	251				300
Rapept-1	LSMKAAFGKV	WKLEENGSSI	IGGAFKAIQA	KNKAPKTTRD	PRLPKPKGQT
Arabpt-1	LSMKAAFGKV	WKLEQNGSSI	IGGTFFKAIQE	RKNAPKAERD	PRLPKPQGQT
Sorghumpt-1	LSMKAAFGKV	WRLEEAGSSI	IGGTIKTIQE	RGKNPKPPRD	PRLPKPKGQT
Mzpt-1	LSMKAAFGKV	WRLEETGSSI	IGGTIKTIQE	RSKNPKPPRD	ARLPKPKGQT
Wtpt-1	LSMKAAFGKV	WRLEEIGSSI	IGGTIKAIQD	KGKNPKPPRD	PRLPAPKGQT
Ricept-1	RALKAAFGKV	WRLEDTGSSI	IGGTIKTIQE	RGKNPKPPRD	PRLPTPKGQT
Cottonpt-1	LSMKAAFGKV	WKLEEIGSSI	IGGTFFKTIQE	RNKTPKPPRD	PRLPKPKGQT
Soybeanpt1	LSMKAAFGKV	WKLEKNNGSSI	IGGTFFKAIQE	RNGASKPPRD	PRLPKPKGQT
Sugpt-1	LSMKAAFGKV	WKLEQKGGSI	IGGTLLKAIQE	RGSNPKPPRD	QRLPKPKGQT
Scpt-1

	301		350
Rapept-1	VGSFRKGLTM LPEAISARLG DKVKVSWKLS SITKLASGEY SLTYETPEGI		
Arabpt-1	VGSFRKGLRM LPEAISARLG SKVKLSWKLS GITKLESGGY NLTYETPDGL		
Sorghumpt-1	VASFRKGLAM LPNAITSSLG SKVKLSWKLT SMTKSDGKGY VLEYETPEGV		
Mzpt-1	VASFRKGLAM LPNAITSSLG SKVKLSWKLT SITKSDDKGY VLEYETPEGV		
Wtpt-1	VASFRKGLAM LPNAIASRLG SKVKLSWKLT SITKADNQGY VLGYETPEGL		
Ricept-1	VASFRKGLTM LPDAITSRLG SKVKLSWKLT SITKSDNKG Y ALVYETPEGV		
Cottonpt-1	VGSFRKGLTM LPEAIANS LG SNVKLSWKLS SITKLGNGGY NLTFETPEGM		
Soybeanpt1	VGSFRKGLTM LPDAISARLG NKVKLSWKLS SISKLDSGEY SLTYETPEGV		
Sugpt-1	VGSFRKGLVM LPTAISARLG SRVKLSWTLS SIVKSLNGEY SLTYDTPDGL		
Scpt-1		
	351		400
Rapept-1	VTVQSKSVVM TVPSHVASSL LRPLSDSAAE ALSKLYYPPV AAVSISYAKE		
Arabpt-1	VSVQSKSVVM TVPSHVASGL LRPLSESAAN ALSKLYYPPV AAVSISYPKE		
Sorghumpt-1	VLVQAKSVIM TIPS YVASDI LRPLSGDAAD VLSRFYYPPV AAVTVSYPKE		
Mzpt-1	VSVQAKSVIM TIPS YVASNI LRPLSSDAAD ALSRFYYPPV AAVTVSYPKE		
Wtpt-1	VSVQAKSVIM TIPS YVASDI LRPLSIDAAD ALSKFYYPPV AAVTVSYPKE		
Ricept-1	VSVQAKTVVM TIPS YVASDI LRPLSSDAAD ALSIFYYPPV AAVTVSYPKE		
Cottonpt-1	VSLQSRSVVM TIPSHVASNL LHPLSAAAAD ALSQFYYPPV ASVTVSYPKE		
Soybeanpt1	VSLQCKTVVL TIPS YVASTL LRPLSAAAAD ALSKFYYPPV AAVSISYPKE		
Sugpt-1	VSVRTKSVVM TVPSYVASRL LRPLSDSAAE SLSKFYYPPV AAVSLSYPKE		
Scpt-1		
	401		450
Rapept-1	AIRSECLIDG ELKGFGQLHP RTQKVETLGT IYSSSLFPNR APPGRVLLLN		
Arabpt-1	AIRTECLIDG ELKGFGQLHP RTQGVETLGT IYSSSLFPNR APPGRILLN		
Sorghumpt-1	AIRKECLIDG ELQGFQQLHP RSQGVETLGT IYSSSLFPNR APAGRVLLLN		
Mzpt-1	AIRKECLIDG ELQGFQQLHP RSQGVETLGT IYSSSLFPNR APDGRVLLLN		
Wtpt-1	AIRKECLIDG ELQGFQQLHP RSQGVETLGT IYSSSLFPNR APAGRVLLLN		
Ricept-1	AIRKECLIDG ELQGFQQLHP RSQGVETLGT IYSSSLFPNR APAGRVLLLN		
Cottonpt-1	AIRKECLIDG ELKGFGQLHP RSQGIETLGT IYSSSLFPNR APSGRVLLLN		
Soybeanpt1	AIRSECLIDG ELKGFGQLHP RSQGVETLGT IYSSSLFPNR APPGRVLLLN		
Sugpt-1	AIRSECLING ELQGFQQLHP RSQGVETLGT IYSSSLFPGR APPGRILILS		
Scpt-1		
	451		500
Rapept-1	YIGGATNTGI LSKSEGELVE AVDRDLRKML IKPSSTDPLV LGVKLWPQAI		
Arabpt-1	YIGGSTNTGI LSKSEGELVE AVDRDLRKML IKPNSTDPLK LGVRVWPQAI		
Sorghumpt-1	YIGGATNTGI VSKTESELVE AVDRDLRKML INPTAVDPLV LGVRVWPQAI		
Mzpt-1	YIGGATNTGI VSKTESELVE AVDRDLRKML INSTAVDPLV LGVRVWPQAI		
Wtpt-1	YIGGSTNTGI VSKTESDLVG AVDRDLRKML INPRAADPLA LGVRVWPQAI		
Ricept-1	YIGGSTNTGI VSKTESELVE AVDRDLRKML INPRAVDPLV LGVRVWPQAI		
Cottonpt-1	YIGGATNTGI LSKTEGELVE AVDRDLRKML INPNAKDPLV LGVRVWPQAI		
Soybeanpt1	YIGGATNTGI LSKTDSELVE TVDRDLRKIL INPNAQDPFV VGVRLWPQAI		
Sugpt-1	YIGGAKNPGI LNKSKDELAK TVDKDLRML INPDAKLPRV LGVRVWPQAI		
Scpt-1 SKTESELVE AVDRDLRKML INPTAVDPLV LGVRVWPQAI		

501

Rapept-1	PQFLIGHIDL	VDAAKASLSS	SGHEGLFLGG	NYVAGVALGR	CVEGAYETAT	550
Arabpt-1	PQFLVGHFID	LDTAKSSLTS	SGYEGLFLGG	NYVAGVALGR	CVEGAYETAI	
Sorghumpt-1	PQFLVGHLDL	LEAAKSALDQ	GGYNGFLFLGG	NYVAGVALGR	CIEGAYESAA	
Mzpt-1	PQFLVGHLDL	LEAAKAALDR	GGYDGLFLGG	NYVAGVALGR	CVEGAYESAS	
Wtpt-1	PQFLIGHLDR	LAAAKSALGQ	GGYDGLFLGG	KYVAGVALGR	CIEGAYESAS	
Ricept-1	PQFLIGHLDH	LEAAKSALGK	GGYDGLFLGG	NYVAGVALGR	CVEGAYESAS	
Cottonpt-1	PQFLVGHLDL	LDSAKMALRD	SGFHGLFLGG	NYVSGVALGR	CVEGAYEVAA	
Soybeanpt1	PQFLVGHLDL	LDVAKASIRN	TGFEGFLFLGG	NYVSGVALGR	CVEGAYEVAA	
Sugpt-1	PQFSIGHFDL	LDAAKAALTD	TGVKGLFLGG	NYVSGVALGR	CIEGAYESAA	
Scpt-1	PQFLVGHLDL	LEAAKSALDR	GGYDGLFLGG	NYVAGVALGR	CVEGAYESAS	

	551	563
Rapept-1	QVNDFMSRYA	YK*
Arabpt-1	EVNNFMSRYA	YK*
Sorghumpt-1	QIYDFLTKYA	YK*
Mzpt-1	QISDFLTKYA	YK*
Wtpt-1	QVSDFLTKYA	YK*
Ricept-1	QISDYLTKYA	YK*
Cottonpt-1	EVKEFLSQYA	YK*
Soybeanpt1	EVNDFLTNRV	YK*
Sugpt-1	EVVDFLSQYS	DK*
Scpt-1	QIYDFLTKYA	YK*

Table 1B

Sub-sequences of herbicide-tolerant protox enzymes comprising point mutations.

Sub-sequence #	Sub-sequence AA sequence	Δ_n AA wild-type	Δ_n AA substitutions	Corresponding AA position of Δ_n in Table 1A	Exemplary mutants
1	AP Δ_1 F	R	C, L	169	Mz88Cys Mz88Leu
2	F Δ_2 S	C	F, L, K	240	Mz159Phe Mz159Leu Mz159Lys
3	Y Δ_3 G	A	V, T, L, C, I	245	pAraC-1Val pAraC-1Thr pAraC-1Leu pAraC-1Cys pAraC-1Ile pMzC-1Val pMzC-1Thr pMzC-1Leu pWhtC-1Val pWhtC-1Thr pSoyC-1Thr pSoyC-1Leu

4	A Δ_4 D	G	S, L	246	pAraC-3Ser pMzC-3Ser pMzC-3Leu pWhtC-3Ser
5	Y Δ_5 P	P	S, H	388	Soy369Ser Soy369His Cot365Ser
6	P Δ_6 A	V	L	390	Wht356Leu
7	Δ_7 IG	Y	C, I, L, T, M, V, A, H, R	451	pAraC-2Cys pAraC-2Ile pAraC-2Leu pAraC-2Thr pAraC-2Met pAraC-2Val pAraC-2Ala pMzC-2Ile pMzC-2Met pSoyC-2Leu pSoyC-2Ile pSugC-2Cys pSugC-2Leu pSugC-2Ile pSugC-2Val pSugC-2Met pCotC-2Cys pCotC-2His pCotC-2Arg
8	YIGG Δ_8	A, S	P	455	Wht421Pro
9	A Δ_9 P	I	T, H, G, N	500	Mz419Thr Mz419His Mz419Gly Mz419Asn Wht466Thr
10	G Δ_{10} A	V	A	536	Wht502Ala Soy517Ala
18	KA Δ_{18} F	A	T, V	256	Mz175Thr Mz175Val
19	Q Δ_{19} H	L	S	418	Mz337Ser
Second-site mutations					
11	Q Δ_{11} S	P	L	143	AraC118Leu
12	IGG Δ_{12}	T	I, A	274	AraC249Ile AraC249Ala
13	SWXL Δ_{13}	S, T	L	330	AraC305Leu
14	L Δ_{14} Y	N	S	450	AraC425Ser
15	G Δ_{15} XGL	Y, H, F, V	C	523	AraC498Cys

Double mutation					
16	T Δ_{16} G	L	S	428	Mz347Ser-453Thr
17	YV Δ_{17} G	A, (S)	T	534	

Table 2

Comparison of the *Arabidopsis* (SEQ ID NO:4) and maize (SEQ ID NO:8) protox-2 amino acid sequences. Identical residues are denoted by the vertical bars between the two sequences. Alignment was performed using the GAP program described in Deveraux *et al.*, *Nucleic Acids Res.* 12:387-395 (1984). Percent similarity: 75.889 / percent identity: 57.905.

```

1 .....MASGAVAD.HQIEAVSGKRVAV 21
1 MLALTASASSASSHPYRHASAHTRRPRLRAVLAMAGSDDPRAAPARSVAV 50
22 VGAGVSGLAAYKLSRGLNVTVFADGRVGGKLSVMQNGLIWDEGANT 71
51 VGAGVSGLAAYRLRQSGVNVTVFEAADRAGGKIRTNSEGGFVWDEGANT 100
72 MTEAEPEVGSLLDDLGLREKQQFPISQKKRYIVRNGVPVMLPTNPIELVT 121
101 MTEGEWEASRLIDDLGLQDKQQYPNSQHKRYIVKDGAPALIPSDPISLMK 150
122 SSVLSTQSKFQILLEPFLWKK....KSSKVS DASA EESVSEFFQRHFGQE 167
151 SSVLSTKSKIALFFEPFLYKKANTRNSGKVSEEHLSESVGSFCERHFGRE 200
168 VVDYLIDPFVGGTSAADPDSLMSKHSFPDLWNVEKSFGSIIVGAIRTKFA 217
201 VVDYFVDPFVAGTSAGDPESLSIRHAFPALWNLERKYGSVIVGAILSKLA 250
218 AKGGKSRDTKSSPGTKKSGRGSFSFKGGMQILPDTLCKSLSHDEINLDSK 267
251 AKGDPVKTRHDSSGKRRNRRVSFSFHGGMQSLINALHNEVGDDNVKLGTE 300
268 VLSLS..YNSGSRQENWSLSCVSHNETQRQ...NPHYDAVIMTAPLCNVK 312
301 VLSLACTFDGVPALGRWSISVDSKDSGDKDLASNQTFDAVIMTAPLSNVR 350
313 EMKVMKGGQPFQNLNPLPEINYMPLSVLITFTTKEKVKRPLEGFGVLIPSK 362
351 RMKFTKGGAPVVLDFLPKMDYLPLSLMVTAFKKDDVKKPLEGFGVLIPYK 400
363 E.QKHGFKTLGTLFSSMMFPDRSPSDVHLYTTFFIGGSRNQELAKASTDEL 411
401 EQQKHGLKTLGTLFSSMMFPDRAPDDQYLYTTFFVGGSHNRDLAGAPTSIL 450

```

```

412 KQVVTSDLQRLLGVEGEPVSVNHYYWRKAFPLYDSSYDSVMEADKMEND 461
    ||:|||||.::|||||:|. |. | | .|||||: .|.||:||||:|.|.:
451 KQLVTSDLKKLLGVEGQPTFVKHVVWGNAFPLYGHDYSSVLEAIEKMEKN 500

462 LPGFFYAGNHRGGLSVGKSIASGCKAADLVISYLESCSNDKKPNDL* 509
    |||||:|:|:|.||. ||||:|||||.||||| .....:
501 LPGFFYAGNSKDGLAVGSVIASGSKAADLAISYLESHTKHNNSH*... 545

```

Table 3A

Cross tolerance of plant protox mutants to various protox inhibitors.

Formula	AraC-1Val	AraC-2Cys	AraC-1Thr	AraC-3Thr	MzC-1Val
XVII	+	+	+	+	+
VIIa	+	+	+	-	+
IV	++	-	++	++	-
XV	+	+	+	+	+
XI	-	+	+	++	+
XVI	-	-	-	-	+
XII	+	-	++	++	++
XIV	+	-	+	+	+
*X					

+ = 10X or more tolerant than WT

++ = 100X or more tolerant than WT

- = no cross tolerance

* = this compound was tested but provided no information

Table 3B

Cross tolerance of plant protox mutants to various protox inhibitors.

Formula	AraC-1Leu	AraC-2Ile	AraC-1Leu + AraC-2Met	AraC-1Leu + AraC-2Leu	AraC-2Ile + AraC-305Leu	AraC-2Cys + AraC-425Ser	AraC-2Leu + AraC-425Ser	AraC-2Met + AraC-425Ser
XVII	+	+	+	+	+	+	+	+
VIIa	++	++	++	++	++	++	++	++
IV	++	-	+	++	+	-	+	+
XV	++	+++	+++	+++	+++	++	+++	++
XI	++	++	++	++	++	++	++	++
XVI	+++	+++	+++	+++	+++	+	++	++
XII								
XIV	++	++	++	++	++	-	++	++

Table 4

Cross tolerance to various protox inhibitors in a seed germination assay.

Formula	Common name	Tolerance
II	acifluorfen	+
III	fomasafen	+
IV	fluoroglycofen	±
IVb	bifenox	+
IVc	oxyfluorfen	+
IVd	lactofen	±
VIIa	fluthiacet-methyl	++
X	sulfentrazone	+
XI	flupropazil	++
XIV	flumiclorac	+
XVI	flumioxazin	+++
XVII		++
XXIa	BAY 11340	+
XXII		++

± ≤ 10X more tolerant than wt

+ ≥ 10X more tolerant than wt

++ ≥ 100X more tolerant than wt

+++ ≥ 1000X more tolerant than wt

Various modifications of the invention described herein will become apparent to those skilled in the art. Such modifications are intended to fall within the scope of the appended claims.

What Is Claimed Is:

1. An isolated DNA molecule comprising a coding sequence that encodes the amino acid sequence set forth in SEQ ID NO:45.
2. The DNA molecule according to claim 1, wherein said DNA molecule comprises the coding sequence set forth in SEQ ID NO:44.
3. A chimeric gene comprising a promoter active in a plant operatively linked to the DNA molecule according to claim 1.
4. The chimeric gene according to claim 3, additionally comprising a signal sequence operatively linked to said DNA molecule, wherein said signal sequence targets the amino acid sequence encoded by said coding sequence to a plastid.
5. The chimeric gene according to claim 3, additionally comprising a signal sequence operatively linked to said DNA molecule, wherein said signal sequence targets the amino acid sequence encoded by said coding sequence to a mitochondria.
6. A recombinant vector comprising the chimeric gene according to claim 3.
7. A host cell comprising the chimeric gene according to claim 3.
8. The host cell according to claim 7, wherein said host cell is selected from the group consisting of a plant cell, a bacterial cell, a yeast cell, and an insect cell.
9. The host cell according to claim 8, wherein said host cell is a plant cell.
10. A chimeric gene comprising a promoter functional in a plant plastid operatively linked to the DNA molecule according to claim 1.
11. The chimeric gene according to claim 10, wherein said plant plastid promoter is a *clpP* gene promoter.

12. The chimeric gene according to claim 10, further comprising a 5' untranslated sequence (5'UTR) from said plastid promoter and a plastid gene 3' untranslated sequence (3' UTR) operatively linked to said DNA molecule.
13. The chimeric gene according to claim 12, wherein said plant plastid promoter is a *clpP* gene promoter, and wherein said 3' UTR is a plastid *rps16* gene 3' untranslated sequence.
14. A plastid transformation vector comprising the chimeric gene according to claim 12.
15. A plant plastid comprising the plastid transformation vector according to claim 14.
16. A plant comprising the plant cell according to claim 9.
17. Seed from the plant according to claim 16.
18. A plant cell comprising the plant plastid according to claim 15.
19. A plant comprising the plant cell according to claim 18.
20. Seed from the plant according to claim 19.
21. An isolated polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 45, wherein said polypeptide has protoporphyrinogen oxidase (protox) activity.
22. A method comprising:
 - (a) obtaining a host cell comprising a DNA molecule according to claim 1; and
 - (b) expressing a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:45 in said host cell.
23. The method according to claim 22, wherein said host cell is a bacterial cell, a yeast cell or an insect cell.
24. The method according to claim 23, wherein said host cell is *E. coli*.

25. A method for assaying a chemical for the ability to inhibit the activity of the polypeptide according to claim 21 comprising:

(a) combining said polypeptide and protoporphyrinogen IX in a first reaction mixture under conditions in which said polypeptide catalyzes the conversion of said protoporphyrinogen IX to protoporphyrin IX;

(b) combining a chemical, said polypeptide and protoporphyrinogen IX in a second reaction mixture under the conditions of step (a);

(c) measuring the fluorescence emission of said first and said second reaction mixtures at about 622 to about 635 nM after excitation of said first and said second reaction mixtures at about 395 to about 410 nM;

(d) comparing the fluorescence emission of said first and said second reaction mixtures at about 622 to about 635 nM;

(e) selecting a chemical that inhibits the activity of said polypeptide if the fluorescence emission of said second reaction mixture is less than the fluorescence emission of said first reaction mixture and if the decrease in fluorescence emission between said first and second reaction mixtures is larger than the margin of error inherent in the measurement of said fluorescence emission.

SEQUENCE LISTING

<110> Syngenta Participations AG

<120> PROTOPORPHYRINOGEN OXIDASE ("PROTOX") GENES

<130> PB/5-31284A

<140>

<141>

<150> US 09/525,152

<151> 2000-03-14

<160> 47

<170> PatentIn Ver. 2.1

<210> 1

<211> 1719

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (31)..(1644)

<223> Arabidopsis protox-1

<400> 1

```

tgacaaaatt ccgaattctc tgcgatttcc atg gag tta tct ctt ctc cgt ccg 54
                               Met Glu Leu Ser Leu Leu Arg Pro
                               1                               5

acg act caa tcg ctt ctt ccg tcg ttt tcg aag ccc aat ctc cga tta 102
Thr Thr Gln Ser Leu Leu Pro Ser Phe Ser Lys Pro Asn Leu Arg Leu
    10                               15                               20

aat gtt tat aag cct ctt aga ctc cgt tgt tca gtg gcc ggt gga cca 150
Asn Val Tyr Lys Pro Leu Arg Leu Arg Cys Ser Val Ala Gly Gly Pro
    25                               30                               35                               40

acc gtc gga tct tca aaa atc gaa ggc gga gga ggc acc acc atc acg 198
Thr Val Gly Ser Ser Lys Ile Glu Gly Gly Gly Gly Thr Thr Ile Thr
                               45                               50                               55

acg gat tgt gtg att gtc ggc gga ggt att agt ggt ctt tgc atc gct 246
Thr Asp Cys Val Ile Val Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala
                               60                               65                               70

cag gcg ctt gct act aag cat cct gat gct gct ccg aat tta att gtg 294
Gln Ala Leu Ala Thr Lys His Pro Asp Ala Ala Pro Asn Leu Ile Val
    75                               80                               85

acc gag gct aag gat cgt gtt gga ggc aac att atc act cgt gaa gag 342
Thr Glu Ala Lys Asp Arg Val Gly Gly Asn Ile Ile Thr Arg Glu Glu
    90                               95                               100

```

aat ggt ttt ctc tgg gaa gaa ggt ccc aat agt ttt caa ccg tct gat	390
Asn Gly Phe Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp	
105 110 115 120	
cct atg ctc act atg gtg gta gat agt ggt ttg aag gat gat ttg gtg	438
Pro Met Leu Thr Met Val Val Asp Ser Gly Leu Lys Asp Asp Leu Val	
125 130 135	
ttg gga gat cct act gcg cca agg ttt gtg ttg tgg aat ggg aaa ttg	486
Leu Gly Asp Pro Thr Ala Pro Arg Phe Val Leu Trp Asn Gly Lys Leu	
140 145 150	
agg ccg gtt cca tcg aag cta aca gac tta ccg ttc ttt gat ttg atg	534
Arg Pro Val Pro Ser Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met	
155 160 165	
agt att ggt ggg aag att aga gct ggt ttt ggt gca ctt ggc att cga	582
Ser Ile Gly Gly Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg	
170 175 180	
ccg tca cct cca ggt cgt gaa gaa tct gtg gag gag ttt gta cgg cgt	630
Pro Ser Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg	
185 190 195 200	
aaac ctc ggt gat gag gtt ttt gag cgc ctg att gaa ccg ttt tgt tca	678
Asn Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser	
205 210 215	
ggg gtt tat gct ggt gat cct tca aaa ctg agc atg aaa gca gcg ttt	726
Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe	
220 225 230	
ggg aag gtt tgg aaa cta gag caa aat ggt gga agc ata ata ggt ggt	774
Gly Lys Val Trp Lys Leu Glu Gln Asn Gly Gly Ser Ile Ile Gly Gly	
235 240 245	
act ttt aag gca att cag gag agg aaa aac gct ccc aag gca gaa cga	822
Thr Phe Lys Ala Ile Gln Glu Arg Lys Asn Ala Pro Lys Ala Glu Arg	
250 255 260	
gac ccg cgc ctg cca aaa cca cag ggc caa aca gtt ggt tct ttc agg	870
Asp Pro Arg Leu Pro Lys Pro Gln Gly Gln Thr Val Gly Ser Phe Arg	
265 270 275 280	
aag gga ctt cga atg ttg cca gaa gca ata tct gca aga tta ggt agc	918
Lys Gly Leu Arg Met Leu Pro Glu Ala Ile Ser Ala Arg Leu Gly Ser	
285 290 295	
aaa gtt aag ttg tct tgg aag ctc tca ggt atc act aag ctg gag agc	966
Lys Val Lys Leu Ser Trp Lys Leu Ser Gly Ile Thr Lys Leu Glu Ser	
300 305 310	
gga gga tac aac tta aca tat gag act cca gat ggt tta gtt tcc gtg	1014
Gly Gly Tyr Asn Leu Thr Tyr Glu Thr Pro Asp Gly Leu Val Ser Val	
315 320 325	

cag agc aaa agt gtt gta atg acg gtg cca tct cat gtt gca agt ggt Gln Ser Lys Ser Val Val Met Thr Val Pro Ser His Val Ala Ser Gly 330 335 340	1062
ctc ttg cgc cct ctt tct gaa tct gct gca aat gca ctc tca aaa cta Leu Leu Arg Pro Leu Ser Glu Ser Ala Ala Asn Ala Leu Ser Lys Leu 345 350 355 360	1110
tat tac cca cca gtt gca gca gta tct atc tcg tac ccg aaa gaa gca Tyr Tyr Pro Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala 365 370 375	1158
atc cga aca gaa tgt ttg ata gat ggt gaa cta aag ggt ttt ggg caa Ile Arg Thr Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln 380 385 390	1206
ttg cat cca cgc acg caa gga gtt gaa aca tta gga act atc tac agc Leu His Pro Arg Thr Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser 395 400 405	1254
tcc tca ctc ttt cca aat cgc gca ccg ccc gga aga att ttg ctg ttg Ser Ser Leu Phe Pro Asn Arg Ala Pro Pro Gly Arg Ile Leu Leu Leu 410 415 420	1302
aac tac att ggc ggg tct aca aac acc gga att ctg tcc aag tct gaa Asn Tyr Ile Gly Gly Ser Thr Asn Thr Gly Ile Leu Ser Lys Ser Glu 425 430 435 440	1350
ggt gag tta gtg gaa gca gtt gac aga gat ttg agg aaa atg cta att Gly Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile 445 450 455	1398
aag cct aat tcg acc gat cca ctt aaa tta gga gtt agg gta tgg cct Lys Pro Asn Ser Thr Asp Pro Leu Lys Leu Gly Val Arg Val Trp Pro 460 465 470	1446
caa gcc att cct cag ttt cta gtt ggt cac ttt gat atc ctt gac acg Gln Ala Ile Pro Gln Phe Leu Val Gly His Phe Asp Ile Leu Asp Thr 475 480 485	1494
gct aaa tca tct cta acg tct tcg ggc tac gaa ggg cta ttt ttg ggt Ala Lys Ser Ser Leu Thr Ser Ser Gly Tyr Glu Gly Leu Phe Leu Gly 490 495 500	1542
ggc aat tac gtc gct ggt gta gcc tta ggc cgg tgt gta gaa ggc gca Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala 505 510 515 520	1590
tat gaa acc gcg att gag gtc aac aac ttc atg tca cgg tac gct tac Tyr Glu Thr Ala Ile Glu Val Asn Asn Phe Met Ser Arg Tyr Ala Tyr 525 530 535	1638
aag taa atgtaaaaca ttaaattctcc cagcttgctg gagtttttatt aaatattttg Lys	1694
agatatccaa aaaaaaaaaa aaaaa	1719

<210> 2
 <211> 537
 <212> PRT
 <213> Arabidopsis thaliana

<400> 2
 Met Glu Leu Ser Leu Leu Arg Pro Thr Thr Gln Ser Leu Leu Pro Ser
 1 5 10 15
 Phe Ser Lys Pro Asn Leu Arg Leu Asn Val Tyr Lys Pro Leu Arg Leu
 20 25 30
 Arg Cys Ser Val Ala Gly Gly Pro Thr Val Gly Ser Ser Lys Ile Glu
 35 40 45
 Gly Gly Gly Gly Thr Thr Ile Thr Thr Asp Cys Val Ile Val Gly Gly
 50 55 60
 Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala Thr Lys His Pro
 65 70 75 80
 Asp Ala Ala Pro Asn Leu Ile Val Thr Glu Ala Lys Asp Arg Val Gly
 85 90 95
 Gly Asn Ile Ile Thr Arg Glu Glu Asn Gly Phe Leu Trp Glu Glu Gly
 100 105 110
 Pro Asn Ser Phe Gln Pro Ser Asp Pro Met Leu Thr Met Val Val Asp
 115 120 125
 Ser Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Thr Ala Pro Arg
 130 135 140
 Phe Val Leu Trp Asn Gly Lys Leu Arg Pro Val Pro Ser Lys Leu Thr
 145 150 155 160
 Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Gly Gly Lys Ile Arg Ala
 165 170 175
 Gly Phe Gly Ala Leu Gly Ile Arg Pro Ser Pro Pro Gly Arg Glu Glu
 180 185 190
 Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly Asp Glu Val Phe Glu
 195 200 205
 Arg Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser
 210 215 220
 Lys Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Lys Leu Glu Gln
 225 230 235 240
 Asn Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys Ala Ile Gln Glu Arg
 245 250 255
 Lys Asn Ala Pro Lys Ala Glu Arg Asp Pro Arg Leu Pro Lys Pro Gln
 260 265 270
 Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Arg Met Leu Pro Glu
 275 280 285
 Ala Ile Ser Ala Arg Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu
 290 295 300
 Ser Gly Ile Thr Lys Leu Glu Ser Gly Gly Tyr Asn Leu Thr Tyr Glu
 305 310 315 320
 Thr Pro Asp Gly Leu Val Ser Val Gln Ser Lys Ser Val Val Met Thr
 325 330 335
 Val Pro Ser His Val Ala Ser Gly Leu Arg Pro Leu Ser Glu Ser
 340 345 350
 Ala Ala Asn Ala Leu Ser Lys Leu Tyr Tyr Pro Pro Val Ala Ala Val
 355 360 365
 Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg Thr Glu Cys Leu Ile Asp
 370 375 380
 Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Thr Gln Gly Val
 385 390 395 400

Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala
 405 410 415
 Pro Pro Gly Arg Ile Leu Leu Leu Asn Tyr Ile Gly Gly Ser Thr Asn
 420 425 430
 Thr Gly Ile Leu Ser Lys Ser Glu Gly Glu Leu Val Glu Ala Val Asp
 435 440 445
 Arg Asp Leu Arg Lys Met Leu Ile Lys Pro Asn Ser Thr Asp Pro Leu
 450 455 460
 Lys Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val
 465 470 475 480
 Gly His Phe Asp Ile Leu Asp Thr Ala Lys Ser Ser Leu Thr Ser Ser
 485 490 495
 Gly Tyr Glu Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala
 500 505 510
 Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Thr Ala Ile Glu Val Asn
 515 520 525
 Asn Phe Met Ser Arg Tyr Ala Tyr Lys
 530 535

<210> 3

<211> 1738

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (70)..(1596)

<223> Arabidopsis protox-2

<400> 3

ttttttacttt atttccgtca ctgcttttcga ctggtcagag attttgactc tgaattgttg 60

cagatagca atg gcg tct gga gca gta gca gat cat caa att gaa gcg gtt 111

Met Ala Ser Gly Ala Val Ala Asp His Gln Ile Glu Ala Val
 1 5 10

tca gga aaa aga gtc gca gtc gta ggt gca ggt gta agt gga ctt gcg 159

Ser Gly Lys Arg Val Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala
 15 20 25 30

gcg gct tac aag ttg aaa tcg agg ggt ttg aat gtg act gtg ttt gaa 207

Ala Ala Tyr Lys Leu Lys Ser Arg Gly Leu Asn Val Thr Val Phe Glu
 35 40 45

gct gat gga aga gta ggt ggg aag ttg aga agt gtt atg caa aat ggt 255

Ala Asp Gly Arg Val Gly Gly Lys Leu Arg Ser Val Met Gln Asn Gly
 50 55 60

ttg att tgg gat gaa gga gca aac acc atg act gag gct gag cca gaa 303

Leu Ile Trp Asp Glu Gly Ala Asn Thr Met Thr Glu Ala Glu Pro Glu
 65 70 75

gtt ggg agt tta ctt gat gat ctt ggg ctt cgt gag aaa caa caa ttt 351

Val Gly Ser Leu Leu Asp Asp Leu Gly Leu Arg Glu Lys Gln Gln Phe
 80 85 90

cca att tca cag aaa aag cgg tat att gtg cgg aat ggt gta cct gtg	399
Pro Ile Ser Gln Lys Lys Arg Tyr Ile Val Arg Asn Gly Val Pro Val	
95 100 105 110	
atg cta cct acc aat ccc ata gag ctg gtc aca agt agt gtg ctc tct	447
Met Leu Pro Thr Asn Pro Ile Glu Leu Val Thr Ser Ser Val Leu Ser	
115 120 125	
acc caa tct aag ttt caa atc ttg ttg gaa cca ttt tta tgg aag aaa	495
Thr Gln Ser Lys Phe Gln Ile Leu Leu Glu Pro Phe Leu Trp Lys Lys	
130 135 140	
aag tcc tca aaa gtc tca gat gca tct gct gaa gaa agt gta agc gag	543
Lys Ser Ser Lys Val Ser Asp Ala Ser Ala Glu Glu Ser Val Ser Glu	
145 150 155	
ttc ttt caa cgc cat ttt gga caa gag gtt gtt gac tat ctc atc gac	591
Phe Phe Gln Arg His Phe Gly Gln Glu Val Val Asp Tyr Leu Ile Asp	
160 165 170	
cct ttt gtt ggt gga aca agt gct gcg gac cct gat tcc ctt tca atg	639
Pro Phe Val Gly Gly Thr Ser Ala Ala Asp Pro Asp Ser Leu Ser Met	
175 180 185 190	
aaag cat tct ttc cca gat ctc tgg aat gta gag aaa agt ttt ggc tct	687
Lys His Ser Phe Pro Asp Leu Trp Asn Val Glu Lys Ser Phe Gly Ser	
195 200 205	
att ata gtc ggt gca atc aga aca aag ttt gct gct aaa ggt ggt aaa	735
Ile Ile Val Gly Ala Ile Arg Thr Lys Phe Ala Ala Lys Gly Gly Lys	
210 215 220	
agt aga gac aca aag agt tct cct ggc aca aaa aag ggt tgc cgt ggg	783
Ser Arg Asp Thr Lys Ser Ser Pro Gly Thr Lys Lys Gly Ser Arg Gly	
225 230 235	
tca ttc tct ttt aag ggg gga atg cag att ctt cct gat acg ttg tgc	831
Ser Phe Ser Phe Lys Gly Gly Met Gln Ile Leu Pro Asp Thr Leu Cys	
240 245 250	
aaa agt ctc tca cat gat gag atc aat tta gac tcc aag gta ctc tct	879
Lys Ser Leu Ser His Asp Glu Ile Asn Leu Asp Ser Lys Val Leu Ser	
255 260 265 270	
ttg tct tac aat tct gga tca aga cag gag aac tgg tca tta tct tgt	927
Leu Ser Tyr Asn Ser Gly Ser Arg Gln Glu Asn Trp Ser Leu Ser Cys	
275 280 285	
gtt tgc cat aat gaa acg cag aga caa aac ccc cat tat gat gct gta	975
Val Ser His Asn Glu Thr Gln Arg Gln Asn Pro His Tyr Asp Ala Val	
290 295 300	
att atg acg gct cct ctg tgc aat gtg aag gag atg aag gtt atg aaa	1023
Ile Met Thr Ala Pro Leu Cys Asn Val Lys Glu Met Lys Val Met Lys	
305 310 315	

```

gga gga caa ccc ttt cag cta aac ttt ctc ccc gag att aat tac atg 1071
Gly Gly Gln Pro Phe Gln Leu Asn Phe Leu Pro Glu Ile Asn Tyr Met
    320                325                330

ccc ctc tcg gtt tta atc acc aca ttc aca aag gag aaa gta aag aga 1119
Pro Leu Ser Val Leu Ile Thr Thr Phe Thr Lys Glu Lys Val Lys Arg
335                340                345                350

cct ctt gaa ggc ttt ggg gta ctc att cca tct aag gag caa aag cat 1167
Pro Leu Glu Gly Phe Gly Val Leu Ile Pro Ser Lys Glu Gln Lys His
                355                360                365

ggt ttc aaa act cta ggt aca ctt ttt tca tca atg atg ttt cca gat 1215
Gly Phe Lys Thr Leu Gly Thr Leu Phe Ser Ser Met Met Phe Pro Asp
                370                375                380

cgt tcc cct agt gac gtt cat cta tat aca act ttt att ggt ggg agt 1263
Arg Ser Pro Ser Asp Val His Leu Tyr Thr Thr Phe Ile Gly Gly Ser
    385                390                395

agg aac cag gaa cta gcc aaa gct tcc act gac gaa tta aaa caa gtt 1311
Arg Asn Gln Glu Leu Ala Lys Ala Ser Thr Asp Glu Leu Lys Gln Val
    400                405                410

gtg act tct gac ctt cag cga ctg ttg ggg gtt gaa ggt gaa ccc gtg 1359
Val Thr Ser Asp Leu Gln Arg Leu Leu Gly Val Glu Gly Glu Pro Val
415                420                425                430

tct gtc aac cat tac tat tgg agg aaa gca ttc ccg ttg tat gac agc 1407
Ser Val Asn His Tyr Tyr Trp Arg Lys Ala Phe Pro Leu Tyr Asp Ser
                435                440                445

agc tat gac tca gtc atg gaa gca att gac aag atg gag aat gat cta 1455
Ser Tyr Asp Ser Val Met Glu Ala Ile Asp Lys Met Glu Asn Asp Leu
                450                455                460

cct ggg ttc ttc tat gca ggt aat cat cga ggg ggg ctc tct gtt ggg 1503
Pro Gly Phe Phe Tyr Ala Gly Asn His Arg Gly Gly Leu Ser Val Gly
    465                470                475

aaa tca ata gca tca ggt tgc aaa gca gct gac ctt gtg atc tca tac 1551
Lys Ser Ile Ala Ser Gly Cys Lys Ala Ala Asp Leu Val Ile Ser Tyr
    480                485                490

ctg gag tct tgc tca aat gac aag aaa cca aat gac agc tta taa 1596
Leu Glu Ser Cys Ser Asn Asp Lys Lys Pro Asn Asp Ser Leu
495                500                505

cattgtcaag gttcgtccct ttttatcact tactttgttaa acttgtaaaa tgcaacaagc 1656

cgccgtgcga ttagccaaca actcagcaaa acccagattc tcataaggct cactaattcc 1716

agaataaact atttatgtaa aa 1738

```

<210> 4

<211> 508

<212> PRT

<213> Arabidopsis thaliana

<400> 4

```

Met Ala Ser Gly Ala Val Ala Asp His Gln Ile Glu Ala Val Ser Gly
 1           5           10           15
Lys Arg Val Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala Ala Ala
          20           25           30
Tyr Lys Leu Lys Ser Arg Gly Leu Asn Val Thr Val Phe Glu Ala Asp
          35           40           45
Gly Arg Val Gly Gly Lys Leu Arg Ser Val Met Gln Asn Gly Leu Ile
          50           55           60
Trp Asp Glu Gly Ala Asn Thr Met Thr Glu Ala Glu Pro Glu Val Gly
          65           70           75           80
Ser Leu Leu Asp Asp Leu Gly Leu Arg Glu Lys Gln Gln Phe Pro Ile
          85           90           95
Ser Gln Lys Lys Arg Tyr Ile Val Arg Asn Gly Val Pro Val Met Leu
          100          105          110
Pro Thr Asn Pro Ile Glu Leu Val Thr Ser Ser Val Leu Ser Thr Gln
          115          120          125
*Ser Lys Phe Gln Ile Leu Leu Glu Pro Phe Leu Trp Lys Lys Lys Ser
          130          135          140
*Ser Lys Val Ser Asp Ala Ser Ala Glu Glu Ser Val Ser Glu Phe Phe
          145          150          155          160
Gln Arg His Phe Gly Gln Glu Val Val Asp Tyr Leu Ile Asp Pro Phe
          165          170          175
*Val Gly Gly Thr Ser Ala Ala Asp Pro Asp Ser Leu Ser Met Lys His
          180          185          190
*Ser Phe Pro Asp Leu Trp Asn Val Glu Lys Ser Phe Gly Ser Ile Ile
          195          200          205
*Val Gly Ala Ile Arg Thr Lys Phe Ala Ala Lys Gly Gly Lys Ser Arg
          210          215          220
*Asp Thr Lys Ser Ser Pro Gly Thr Lys Lys Gly Ser Arg Gly Ser Phe
          225          230          235          240
*Ser Phe Lys Gly Gly Met Gln Ile Leu Pro Asp Thr Leu Cys Lys Ser
          245          250          255
*Leu Ser His Asp Glu Ile Asn Leu Asp Ser Lys Val Leu Ser Leu Ser
          260          265          270
Tyr Asn Ser Gly Ser Arg Gln Glu Asn Trp Ser Leu Ser Cys Val Ser
          275          280          285
His Asn Glu Thr Gln Arg Gln Asn Pro His Tyr Asp Ala Val Ile Met
          290          295          300
Thr Ala Pro Leu Cys Asn Val Lys Glu Met Lys Val Met Lys Gly Gly
          305          310          315          320
Gln Pro Phe Gln Leu Asn Phe Leu Pro Glu Ile Asn Tyr Met Pro Leu
          325          330          335
Ser Val Leu Ile Thr Thr Phe Thr Lys Glu Lys Val Lys Arg Pro Leu
          340          345          350
Glu Gly Phe Gly Val Leu Ile Pro Ser Lys Glu Gln Lys His Gly Phe
          355          360          365
Lys Thr Leu Gly Thr Leu Phe Ser Ser Met Met Phe Pro Asp Arg Ser
          370          375          380
Pro Ser Asp Val His Leu Tyr Thr Thr Phe Ile Gly Gly Ser Arg Asn
          385          390          395          400
Gln Glu Leu Ala Lys Ala Ser Thr Asp Glu Leu Lys Gln Val Val Thr
          405          410          415
Ser Asp Leu Gln Arg Leu Leu Gly Val Glu Gly Glu Pro Val Ser Val

```

```

          420          425          430
Asn His Tyr Tyr Trp Arg Lys Ala Phe Pro Leu Tyr Asp Ser Ser Tyr
          435          440          445
Asp Ser Val Met Glu Ala Ile Asp Lys Met Glu Asn Asp Leu Pro Gly
          450          455          460
Phe Phe Tyr Ala Gly Asn His Arg Gly Gly Leu Ser Val Gly Lys Ser
465          470          475          480
Ile Ala Ser Gly Cys Lys Ala Ala Asp Leu Val Ile Ser Tyr Leu Glu
          485          490          495
Ser Cys Ser Asn Asp Lys Lys Pro Asn Asp Ser Leu
          500          505

```

<210> 5

<211> 1691

<212> DNA

<213> Zea mays

<220>

<221> CDS

<222> (1)..(1443)

<223> maize protox-1 c-DNA(not full-length)

<400> 5

```

gcg gac tgc gtc gtg gtg ggc gga ggc atc agt ggc ctc tgc acc gcg 48
Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu Cys Thr Ala
  1          5          10          15

```

```

cag gcg ctg gcc acg cgg cac ggc gtc ggg gac gtg ctt gtc acg gag 96
Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu Val Thr Glu
          20          25          30

```

```

gcc cgc gcc cgc ccc ggc ggc aac att acc acc gtc gag cgc ccc gag 144
Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Glu
          35          40          45

```

```

gaa ggg tac ctc tgg gag gag ggt ccc aac agc ttc cag ccc tcc gac 192
Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp
          50          55          60

```

```

ccc gtt ctc acc atg gcc gtg gac agc gga ctg aag gat gac ttg gtt 240
Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val
          65          70          75          80

```

```

ttt ggg gac cca aac gcg ccg cgt ttc gtg ctg tgg gag ggg aag ctg 288
Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu
          85          90          95

```

```

agg ccc gtg cca tcc aag ccc gcc gac ctc ccg ttc ttc gat ctc atg 336
Arg Pro Val Pro Ser Lys Pro Ala Asp Leu Pro Phe Phe Asp Leu Met
          100          105          110

```

```

agc atc cca ggg aag ctc agg gcc ggt cta ggc gcg ctt ggc atc cgc 384
Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg
          115          120          125

```


ccg cct cct cca ggc cgc gaa gag tca gtg gag gag ttc gtg cgc cgc	432
Pro Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg	
130 135 140	
aac ctc ggt gct gag gtc ttt gag cgc ctc att gag cct ttc tgc tca	480
Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser	
145 150 155 160	
ggt gtc tat gct ggt gat cct tct aag ctc agc atg aag gct gca ttt	528
Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe	
165 170 175	
ggg aag gtt tgg cgg ttg gaa gaa act gga ggt agt att att ggt gga	576
Gly Lys Val Trp Arg Leu Glu Glu Thr Gly Gly Ser Ile Ile Gly Gly	
180 185 190	
acc atc aag aca att cag gag agg agc aag aat cca aaa cca ccg agg	624
Thr Ile Lys Thr Ile Gln Glu Arg Ser Lys Asn Pro Lys Pro Pro Arg	
195 200 205	
agat gcc cgc ctt ccg aag cca aaa ggg cag aca gtt gca tct ttc agg	672
Asp Ala Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Ala Ser Phe Arg	
210 215 220	
aag ggt ctt gcc atg ctt cca aat gcc att aca tcc agc ttg ggt agt	720
Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Thr Ser Ser Leu Gly Ser	
225 230 235 240	
aaa gtc aaa cta tca tgg aaa ctc acg agc att aca aaa tca gat gac	768
Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ser Asp Asp	
245 250 255	
aag gga tat gtt ttg gag tat gaa acg cca gaa ggg gtt gtt tcg gtg	816
Lys Gly Tyr Val Leu Glu Tyr Glu Thr Pro Glu Gly Val Val Ser Val	
260 265 270	
cag gct aaa agt gtt atc atg act att cca tca tat gtt gct agc aac	864
Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asn	
275 280 285	
att ttg cgt cca ctt tca agc gat gct gca gat gct cta tca aga ttc	912
Ile Leu Arg Pro Leu Ser Ser Asp Ala Ala Asp Ala Leu Ser Arg Phe	
290 295 300	
tat tat cca ccg gtt gct gct gta act gtt tcg tat cca aag gaa gca	960
Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala	
305 310 315 320	
att aga aaa gaa tgc tta att gat ggg gaa ctc cag ggc ttt ggc cag	1008
Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln	
325 330 335	
ttg cat cca cgt agt caa gga gtt gag aca tta gga aca ata tac agt	1056
Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser	
340 345 350	
tcc tca ctc ttt cca aat cgt gct cct gac ggt agg gtg tta ctt cta	1104

Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp Gly Arg Val Leu Leu Leu
 355 360 365
 aac tac ata gga ggt gct aca aac aca gga att gtt tcc aag act gaa 1152
 Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser Lys Thr Glu
 370 375 380
 agt gag ctg gtc gaa gca gtt gac cgt gac ctc cga aaa atg ctt ata 1200
 Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile
 385 390 395 400
 aat tct aca gca gtg gac cct tta gtc ctt ggt gtt cga gtt tgg cca 1248
 Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg Val Trp Pro
 405 410 415
 caa gcc ata cct cag ttc ctg gta gga cat ctt gat ctt ctg gaa gcc 1296
 Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Glu Ala
 420 425 430
 gca aaa gct gcc ctg gac cga ggt ggc tac gat ggg ctg ttc cta gga 1344
 Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu Phe Leu Gly
 435 440 445
 ggg aac tat gtt gca gga gtt gcc ctg ggc aga tgc gtt gag ggc gcg 1392
 Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala
 450 455 460
 tat gaa agt gcc tcg caa ata tct gac ttc ttg acc aag tat gcc tac 1440
 Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys Tyr Ala Tyr
 465 470 475 480
 aag tgatgaaaga agtggagcgc tacttggttaa tcgtttatgt tgcatagatg 1493
 Lys
 aggtgcctcc ggggaaaaaa aagcttgaat agtatTTTTT attcttattt tgtaaattgc 1553
 atttctgttc ttttttctat cagtaattag ttatatTTTA gttctgtagg agattgttct 1613
 gttcactgcc cttcaaaaga aattttattt ttcattcttt tatgagagct gtgctactta 1673
 aaaaaaaaaa aaaaaaaaaa 1691

<210> 6
 <211> 481
 <212> PRT
 <213> Zea mays

<400> 6
 Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu Cys Thr Ala
 1 5 10 15
 Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu Val Thr Glu
 20 25 30
 Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Glu
 35 40 45

Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp
 50 55 60
 Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val
 65 70 75 80
 Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu
 85 90 95
 Arg Pro Val Pro Ser Lys Pro Ala Asp Leu Pro Phe Phe Asp Leu Met
 100 105 110
 Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg
 115 120 125
 Pro Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg
 130 135 140
 Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser
 145 150 155 160
 Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe
 165 170 175
 Gly Lys Val Trp Arg Leu Glu Glu Thr Gly Gly Ser Ile Ile Gly Gly
 180 185 190
 Thr Ile Lys Thr Ile Gln Glu Arg Ser Lys Asn Pro Lys Pro Pro Arg
 195 200 205
 Asp Ala Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Ala Ser Phe Arg
 210 215 220
 Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Thr Ser Ser Leu Gly Ser
 225 230 235 240
 Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ser Asp Asp
 245 250 255
 Lys Gly Tyr Val Leu Glu Tyr Glu Thr Pro Glu Gly Val Val Ser Val
 260 265 270
 Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asn
 275 280 285
 Ile Leu Arg Pro Leu Ser Ser Asp Ala Ala Asp Ala Leu Ser Arg Phe
 290 295 300
 Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala
 305 310 315 320
 Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln
 325 330 335
 Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser
 340 345 350

Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp Gly Arg Val Leu Leu Leu
 355 360 365

Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser Lys Thr Glu
 370 375 380

Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile
 385 390 395 400

Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg Val Trp Pro
 405 410 415

Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Glu Ala
 420 425 430

Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu Phe Leu Gly
 435 440 445

Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala
 450 455 460

Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys Tyr Ala Tyr
 465 470 475 480

Lys

<210> 7
 <211> 2061
 <212> DNA
 <213> Zea mays

<220>
 <221> CDS
 <222> (64)..(1698)
 <223> maize protox-2

<400> 7
 ctctcctacc tccacctcca cgacaacaag caaatcccca tccagttcca aaccctaact 60

caa atg ctc gct ttg act gcc tca gcc tca tcc gct tcg tcc cat cct 108
 Met Leu Ala Leu Thr Ala Ser Ala Ser Ser Ala Ser Ser His Pro
 1 5 10 15

tat cgc cac gcc tcc gcg cac act cgt cgc ccc cgc cta cgt gcg gtc 156
 Tyr Arg His Ala Ser Ala His Thr Arg Arg Pro Arg Leu Arg Ala Val
 20 25 30

ctc gcg atg gcg ggc tcc gac gac ccc cgt gca gcg ccc gcc aga tcg 204
 Leu Ala Met Ala Gly Ser Asp Asp Pro Arg Ala Ala Pro Ala Arg Ser
 35 40 45

gtc gcc gtc gtc ggc gcc ggg gtc agc ggg ctc gcg gcg gcg tac agg 252
 Val Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala Ala Ala Tyr Arg

50	55	60	
ctc aga cag agc ggc gtg aac gta acg gtg ttc gaa gcg gcc gac agg Leu Arg Gln Ser Gly Val Asn Val Thr Val Phe Glu Ala Ala Asp Arg 65 70 75			300
gcg gga gga aag ata cgg acc aat tcc gag ggc ggg ttt gtc tgg gat Ala Gly Gly Lys Ile Arg Thr Asn Ser Glu Gly Gly Phe Val Trp Asp 80 85 90 95			348
gaa gga gct aac acc atg aca gaa ggt gaa tgg gag gcc agt aga ctg Glu Gly Ala Asn Thr Met Thr Glu Gly Glu Trp Glu Ala Ser Arg Leu 100 105 110			396
att gat gat ctt ggt cta caa gac aaa cag cag tat cct aac tcc caa Ile Asp Asp Leu Gly Leu Gln Asp Lys Gln Gln Tyr Pro Asn Ser Gln 115 120 125			444
cac aag cgt tac att gtc aaa gat gga gca cca gca ctg att cct tcg His Lys Arg Tyr Ile Val Lys Asp Gly Ala Pro Ala Leu Ile Pro Ser 130 135 140			492
gat ccc att tcg cta atg aaa agc agt gtt ctt tcg aca aaa tca aag Asp Pro Ile Ser Leu Met Lys Ser Ser Val Leu Ser Thr Lys Ser Lys 145 150 155			540
att gcg tta ttt ttt gaa cca ttt ctc tac aag aaa gct aac aca aga Ile Ala Leu Phe Phe Glu Pro Phe Leu Tyr Lys Lys Ala Asn Thr Arg 160 165 170 175			588
aac tct gga aaa gtg tct gag gag cac ttg agt gag agt gtt ggg agc Asn Ser Gly Lys Val Ser Glu Glu His Leu Ser Glu Ser Val Gly Ser 180 185 190			636
ttc tgt gaa cgc cac ttt gga aga gaa gtt gtt gac tat ttt gtt gat Phe Cys Glu Arg His Phe Gly Arg Glu Val Val Asp Tyr Phe Val Asp 195 200 205			684
cca ttt gta gct gga aca agt gca gga gat cca gag tca cta tct att Pro Phe Val Ala Gly Thr Ser Ala Gly Asp Pro Glu Ser Leu Ser Ile 210 215 220			732
cgt cat gca ttc cca gca ttg tgg aat ttg gaa aga aag tat ggt tca Arg His Ala Phe Pro Ala Leu Trp Asn Leu Glu Arg Lys Tyr Gly Ser 225 230 235			780
gtt att gtt ggt gcc atc ttg tct aag cta gca gct aaa ggt gat cca Val Ile Val Gly Ala Ile Leu Ser Lys Leu Ala Ala Lys Gly Asp Pro 240 245 250 255			828
gta aag aca aga cat gat tca tca ggg aaa aga agg aat aga cga gtg Val Lys Thr Arg His Asp Ser Ser Gly Lys Arg Arg Asn Arg Arg Val 260 265 270			876
tcg ttt tca ttt cat ggt gga atg cag tca cta ata aat gca ctt cac Ser Phe Ser Phe His Gly Gly Met Gln Ser Leu Ile Asn Ala Leu His 275 280 285			924

aat gaa gtt gga gat gat aat gtg aag ctt ggt aca gaa gtg ttg tca 972
 Asn Glu Val Gly Asp Asp Asn Val Lys Leu Gly Thr Glu Val Leu Ser
 290 295 300

ttg gca tgt aca ttt gat gga gtt cct gca cta ggc agg tgg tca att 1020
 Leu Ala Cys Thr Phe Asp Gly Val Pro Ala Leu Gly Arg Trp Ser Ile
 305 310 315

tct gtt gat tcg aag gat agc ggt gac aag gac ctt gct agt aac caa 1068
 Ser Val Asp Ser Lys Asp Ser Gly Asp Lys Asp Leu Ala Ser Asn Gln
 320 325 330 335

acc ttt gat gct gtt ata atg aca gct cca ttg tca aat gtc cgg agg 1116
 Thr Phe Asp Ala Val Ile Met Thr Ala Pro Leu Ser Asn Val Arg Arg
 340 345 350

atg aag ttc acc aaa ggt gga gct ccg gtt gtt ctt gac ttt ctt cct 1164
 Met Lys Phe Thr Lys Gly Gly Ala Pro Val Val Leu Asp Phe Leu Pro
 355 360 365

aag atg gat tat cta cca cta tct ctc atg gtg act gct ttt aag aag 1212
 Lys Met Asp Tyr Leu Pro Leu Ser Leu Met Val Thr Ala Phe Lys Lys
 370 375 380

gat gat gtc aag aaa cct ctg gaa gga ttt ggg gtc tta ata cct tac 1260
 Asp Asp Val Lys Lys Pro Leu Glu Gly Phe Gly Val Leu Ile Pro Tyr
 385 390 395

aag gaa cag caa aaa cat ggt ctg aaa acc ctt ggg act ctc ttt tcc 1308
 Lys Glu Gln Gln Lys His Gly Leu Lys Thr Leu Gly Thr Leu Phe Ser
 400 405 410 415

tca atg atg ttc cca gat cga gct cct gat gac caa tat tta tat aca 1356
 Ser Met Met Phe Pro Asp Arg Ala Pro Asp Asp Gln Tyr Leu Tyr Thr
 420 425 430

aca ttt gtt ggg ggt agc cac aat aga gat ctt gct gga gct cca acg 1404
 Thr Phe Val Gly Gly Ser His Asn Arg Asp Leu Ala Gly Ala Pro Thr
 435 440 445

tct att ctg aaa caa ctt gtg acc tct gac ctt aaa aaa ctc ttg ggc 1452
 Ser Ile Leu Lys Gln Leu Val Thr Ser Asp Leu Lys Lys Leu Leu Gly
 450 455 460

gta gag ggg caa cca act ttt gtc aag cat gta tac tgg gga aat gct 1500
 Val Glu Gly Gln Pro Thr Phe Val Lys His Val Tyr Trp Gly Asn Ala
 465 470 475

ttt cct ttg tat ggc cat gat tat agt tct gta ttg gaa gct ata gaa 1548
 Phe Pro Leu Tyr Gly His Asp Tyr Ser Ser Val Leu Glu Ala Ile Glu
 480 485 490 495

aag atg gag aaa aac ctt cca ggg ttc ttc tac gca gga aat agc aag 1596
 Lys Met Glu Lys Asn Leu Pro Gly Phe Phe Tyr Ala Gly Asn Ser Lys
 500 505 510

gat ggg ctt gct gtt gga agt gtt ata gct tca gga agc aag gct gct 1644
 Asp Gly Leu Ala Val Gly Ser Val Ile Ala Ser Gly Ser Lys Ala Ala
 515 520 525
 gac ctt gca atc tca tat ctt gaa tct cac acc aag cat aat aat tca 1692
 Asp Leu Ala Ile Ser Tyr Leu Glu Ser His Thr Lys His Asn Asn Ser
 530 535 540
 cat tga aagtgtctga cctatcctct agcagttgtc gacaaatttc tccagttcat 1748
 His
 545
 gtacagtaga aaccgatgcg ttgcagtttc agaacatctt cacttcttca gatattaacc 1808
 cttcgttgaa catccaccag aaaggtagtc acatgtgtaa gtgggaaaat gaggttaaaa 1868
 actattatgg cggccgaaat gttccttttt gttttcctca caagtggcct acgacacttg 1928
 atgttggaata tacatttaaa tttgttgaat tgtttgagaa cacatgcgtg acgtgtaata 1988
 tttgcctatt gtgatttttag cagtagtctt ggccagatta tgctttacgc ctttaaaaaa 2048
 aaaaaaaaaa aaa 2061

<210> 8

<211> 544

<212> PRT

<213> Zea mays

<400> 8

Met Leu Ala Leu Thr Ala Ser Ala Ser Ser Ala Ser Ser His Pro Tyr
 1 5 10 15
 Arg His Ala Ser Ala His Thr Arg Arg Pro Arg Leu Arg Ala Val Leu
 20 25 30
 Ala Met Ala Gly Ser Asp Asp Pro Arg Ala Ala Pro Ala Arg Ser Val
 35 40 45
 Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala Ala Ala Tyr Arg Leu
 50 55 60
 Arg Gln Ser Gly Val Asn Val Thr Val Phe Glu Ala Ala Asp Arg Ala
 65 70 75 80
 Gly Gly Lys Ile Arg Thr Asn Ser Glu Gly Gly Phe Val Trp Asp Glu
 85 90 95
 Gly Ala Asn Thr Met Thr Glu Gly Glu Trp Glu Ala Ser Arg Leu Ile
 100 105 110
 Asp Asp Leu Gly Leu Gln Asp Lys Gln Gln Tyr Pro Asn Ser Gln His
 115 120 125
 Lys Arg Tyr Ile Val Lys Asp Gly Ala Pro Ala Leu Ile Pro Ser Asp
 130 135 140
 Pro Ile Ser Leu Met Lys Ser Ser Val Leu Ser Thr Lys Ser Lys Ile
 145 150 155 160
 Ala Leu Phe Phe Glu Pro Phe Leu Tyr Lys Lys Ala Asn Thr Arg Asn
 165 170 175
 Ser Gly Lys Val Ser Glu Glu His Leu Ser Glu Ser Val Gly Ser Phe
 180 185 190
 Cys Glu Arg His Phe Gly Arg Glu Val Val Asp Tyr Phe Val Asp Pro
 195 200 205

Phe Val Ala Gly Thr Ser Ala Gly Asp Pro Glu Ser Leu Ser Ile Arg
 210 215 220
 His Ala Phe Pro Ala Leu Trp Asn Leu Glu Arg Lys Tyr Gly Ser Val
 225 230 235 240
 Ile Val Gly Ala Ile Leu Ser Lys Leu Ala Ala Lys Gly Asp Pro Val
 245 250 255
 Lys Thr Arg His Asp Ser Ser Gly Lys Arg Arg Asn Arg Arg Val Ser
 260 265 270
 Phe Ser Phe His Gly Gly Met Gln Ser Leu Ile Asn Ala Leu His Asn
 275 280 285
 Glu Val Gly Asp Asp Asn Val Lys Leu Gly Thr Glu Val Leu Ser Leu
 290 295 300
 Ala Cys Thr Phe Asp Gly Val Pro Ala Leu Gly Arg Trp Ser Ile Ser
 305 310 315 320
 Val Asp Ser Lys Asp Ser Gly Asp Lys Asp Leu Ala Ser Asn Gln Thr
 325 330 335
 Phe Asp Ala Val Ile Met Thr Ala Pro Leu Ser Asn Val Arg Arg Met
 340 345 350
 Lys Phe Thr Lys Gly Gly Ala Pro Val Val Leu Asp Phe Leu Pro Lys
 355 360 365
 Met Asp Tyr Leu Pro Leu Ser Leu Met Val Thr Ala Phe Lys Lys Asp
 370 375 380
 Asp Val Lys Lys Pro Leu Glu Gly Phe Gly Val Leu Ile Pro Tyr Lys
 385 390 395 400
 Glu Gln Gln Lys His Gly Leu Lys Thr Leu Gly Thr Leu Phe Ser Ser
 405 410 415
 Met Met Phe Pro Asp Arg Ala Pro Asp Asp Gln Tyr Leu Tyr Thr Thr
 420 425 430
 Phe Val Gly Gly Ser His Asn Arg Asp Leu Ala Gly Ala Pro Thr Ser
 435 440 445
 Ile Leu Lys Gln Leu Val Thr Ser Asp Leu Lys Lys Leu Leu Gly Val
 450 455 460
 Glu Gly Gln Pro Thr Phe Val Lys His Val Tyr Trp Gly Asn Ala Phe
 465 470 475 480
 Pro Leu Tyr Gly His Asp Tyr Ser Ser Val Leu Glu Ala Ile Glu Lys
 485 490 495
 Met Glu Lys Asn Leu Pro Gly Phe Phe Tyr Ala Gly Asn Ser Lys Asp
 500 505 510
 Gly Leu Ala Val Gly Ser Val Ile Ala Ser Gly Ser Lys Ala Ala Asp
 515 520 525
 Leu Ala Ile Ser Tyr Leu Glu Ser His Thr Lys His Asn Asn Ser His
 530 535 540

<210> 9
 <211> 1811
 <212> DNA
 <213> Triticum aestivum

<220>
 <221> CDS
 <222> (3)..(1589)
 <223> wheat protox-1

<400> 9
 gc gca aca atg gcc acc gcc acc gtc gcg gcc gcg tcg ccg ctc cgc

47

Ala	Thr	Met	Ala	Thr	Ala	Thr	Val	Ala	Ala	Ala	Ser	Pro	Leu	Arg		
1				5				10						15		
ggc	agg	gtc	acc	ggg	cgc	cca	cac	cgc	gtc	cgc	ccg	cgt	tgc	gct	acc	95
Gly	Arg	Val	Thr	Gly	Arg	Pro	His	Arg	Val	Arg	Pro	Arg	Cys	Ala	Thr	
				20				25						30		
gcg	agc	agc	gcg	acc	gag	act	ccg	gcg	gcg	ccc	ggc	gtg	cgg	ctg	tcc	143
Ala	Ser	Ser	Ala	Thr	Glu	Thr	Pro	Ala	Ala	Pro	Gly	Val	Arg	Leu	Ser	
			35					40					45			
gcg	gaa	tgc	gtc	att	gtg	ggc	gcc	ggc	atc	agc	ggc	ctc	tgc	acc	gcg	191
Ala	Glu	Cys	Val	Ile	Val	Gly	Ala	Gly	Ile	Ser	Gly	Leu	Cys	Thr	Ala	
		50					55					60				
cag	gcg	ctg	gcc	acc	cga	tac	ggc	gtc	agc	gac	ctg	ctc	gtc	acg	gag	239
Gln	Ala	Leu	Ala	Thr	Arg	Tyr	Gly	Val	Ser	Asp	Leu	Leu	Val	Thr	Glu	
	65					70					75					
gcc	cgc	gac	cgc	ccg	ggc	ggc	aac	atc	acc	acc	gtc	gag	cgt	ccc	gac	287
Ala	Arg	Asp	Arg	Pro	Gly	Gly	Asn	Ile	Thr	Thr	Val	Glu	Arg	Pro	Asp	
	80				85					90					95	
gag	ggg	tac	ctg	tgg	gag	gag	gga	ccc	aac	agc	ttc	cag	ccc	tcc	gac	335
Glu	Gly	Tyr	Leu	Trp	Glu	Glu	Gly	Pro	Asn	Ser	Phe	Gln	Pro	Ser	Asp	
			100					105					110			
ccg	gtc	ctc	acc	atg	gcc	gtg	gac	agc	ggg	ctc	aag	gat	gac	ttg	gtg	383
Pro	Val	Leu	Thr	Met	Ala	Val	Asp	Ser	Gly	Leu	Lys	Asp	Asp	Leu	Val	
			115					120					125			
ttc	ggg	gac	ccc	aac	ggc	ccc	cgg	ttc	gtg	ctg	tgg	gag	ggg	aag	ctg	431
Phe	Gly	Asp	Pro	Asn	Ala	Pro	Arg	Phe	Val	Leu	Trp	Glu	Gly	Lys	Leu	
			130				135					140				
agg	ccg	gtg	ccg	tcg	aag	cca	ggc	gac	ctg	cct	ttc	ttc	agc	ctc	atg	479
Arg	Pro	Val	Pro	Ser	Lys	Pro	Gly	Asp	Leu	Pro	Phe	Phe	Ser	Leu	Met	
			145			150					155					
agt	atc	cct	ggg	aag	ctc	agg	gcc	ggc	ctt	ggc	gcg	ctc	ggc	att	cgc	527
Ser	Ile	Pro	Gly	Lys	Leu	Arg	Ala	Gly	Leu	Gly	Ala	Leu	Gly	Ile	Arg	
	160				165					170					175	
cca	cct	cct	cca	ggg	cgc	gag	gag	tcg	gtg	gag	gag	ttt	gtg	cgc	cgc	575
Pro	Pro	Pro	Pro	Gly	Arg	Glu	Glu	Ser	Val	Glu	Glu	Phe	Val	Arg	Arg	
				180					185					190		
aac	ctc	ggt	gcc	gag	gtc	ttt	gag	cgc	ctc	atc	gag	cct	ttc	tgc	tca	623
Asn	Leu	Gly	Ala	Glu	Val	Phe	Glu	Arg	Leu	Ile	Glu	Pro	Phe	Cys	Ser	
			195					200					205			
ggt	gta	tat	gct	ggt	gat	cct	tcg	aag	ctt	agt	atg	aag	gct	gca	ttt	671
Gly	Val	Tyr	Ala	Gly	Asp	Pro	Ser	Lys	Leu	Ser	Met	Lys	Ala	Ala	Phe	
		210				215						220				
ggg	aag	gtc	tgg	agg	ttg	gag	gag	att	gga	ggt	agt	att	att	ggt	gga	719
Gly	Lys	Val	Trp	Arg	Leu	Glu	Glu	Ile	Gly	Gly	Ser	Ile	Ile	Gly	Gly	

225	230	235	
acc atc aag gcg att cag gat aaa ggg aag aac ccc aaa ccg cca agg			767
Thr Ile Lys Ala Ile Gln Asp Lys Gly Lys Asn Pro Lys Pro Pro Arg			
240	245	250	255
gat ccc cga ctt ccg gca cca aag gga cag acg gtg gca tct ttc agg			815
Asp Pro Arg Leu Pro Ala Pro Lys Gly Gln Thr Val Ala Ser Phe Arg			
260	265	270	
aag ggt cta gcc atg ctc ccg aat gcc atc gca tct agg ctg ggt agt			863
Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Ala Ser Arg Leu Gly Ser			
275	280	285	
aaa gtc aag ctg tca tgg aag ctt acg agc att aca aag gcg gac aac			911
Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ala Asp Asn			
290	295	300	
caa gga tat gta tta ggt tat gaa aca cca gaa gga ctt gtt tca gtg			959
Gln Gly Tyr Val Leu Gly Tyr Glu Thr Pro Glu Gly Leu Val Ser Val			
305	310	315	
cag gct aaa agt gtt atc atg acc atc ccg tca tat gtt gct agt gat			1007
Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asp			
320	325	330	335
atc ttg cgc cca ctt tca att gat gca gca gat gca ctc tca aaa ttc			1055
Ile Leu Arg Pro Leu Ser Ile Asp Ala Ala Asp Ala Leu Ser Lys Phe			
340	345	350	
tat tat ccg cca gtt gct gct gta act gtt tca tat cca aaa gaa gct			1103
Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala			
355	360	365	
att aga aaa gaa tgc tta att gat ggg gag ctc cag ggt ttc ggc cag			1151
Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln			
370	375	380	
ttg cat cca cgt agc caa gga gtc gag act tta ggg aca ata tat agc			1199
Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser			
385	390	395	
tct tct ctc ttt cct aat cgt gct cct gct gga aga gtg tta ctt ctg			1247
Ser Ser Leu Phe Pro Asn Arg Ala Pro Ala Gly Arg Val Leu Leu Leu			
400	405	410	415
aac tat atc ggg ggt tct aca aat aca ggg atc gtc tcc aag act gag			1295
Asn Tyr Ile Gly Gly Ser Thr Asn Thr Gly Ile Val Ser Lys Thr Glu			
420	425	430	
agt gac tta gta gga gcc gtt gac cgt gac ctc aga aaa atg ttg ata			1343
Ser Asp Leu Val Gly Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile			
435	440	445	
aac cct aga gca gca gac cct tta gca tta ggg gtt cga gtg tgg cca			1391
Asn Pro Arg Ala Ala Asp Pro Leu Ala Leu Gly Val Arg Val Trp Pro			
450	455	460	

caa gca ata cca cag ttt ttg att ggg cac ctt gat cgc ctt gct gct 1439
 Gln Ala Ile Pro Gln Phe Leu Ile Gly His Leu Asp Arg Leu Ala Ala
 465 470 475

gca aaa tct gca ctg ggc caa ggc ggc tac gac ggg ttg ttc cta gga 1487
 Ala Lys Ser Ala Leu Gly Gln Gly Gly Tyr Asp Gly Leu Phe Leu Gly
 480 485 490 495

gga aac tac gtc gca gga gtt gcc ttg ggc cga tgc atc gag ggt gcg 1535
 Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala
 500 505 510

tac gag agt gcc tca caa gta tct gac ttc ttg acc aag tat gcc tac 1583
 Tyr Glu Ser Ala Ser Gln Val Ser Asp Phe Leu Thr Lys Tyr Ala Tyr
 515 520 525

aag tga tggaagtagt gcattctcttc attttgttgc atatacgagg tgaggctagg 1639
 Lys

atcggtaaaa catcatgaga ttctgtagtg tttctttaat tgaaaaaaca aatttttagtg 1699

atgcaatatg tgctctttcc ttagttcga gcatgtacat cggtatggga taaagtagaa 1759

taagctattc tgcaaaagca gtgatttttt ttgaaaaaaa aaaaaaaaaa aa 1811

<210> 10

<211> 528

<212> PRT

<213> Triticum aestivum

<400> 10

Ala Thr Met Ala Thr Ala Thr Val Ala Ala Ala Ser Pro Leu Arg Gly
 1 5 10 15

Arg Val Thr Gly Arg Pro His Arg Val Arg Pro Arg Cys Ala Thr Ala
 20 25 30

Ser Ser Ala Thr Glu Thr Pro Ala Ala Pro Gly Val Arg Leu Ser Ala
 35 40 45

Glu Cys Val Ile Val Gly Ala Gly Ile Ser Gly Leu Cys Thr Ala Gln
 50 55 60

Ala Leu Ala Thr Arg Tyr Gly Val Ser Asp Leu Leu Val Thr Glu Ala
 65 70 75 80

Arg Asp Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Asp Glu
 85 90 95

Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro
 100 105 110

Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val Phe
 115 120 125

Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu Arg
 130 135 140

Pro Val Pro Ser Lys Pro Gly Asp Leu Pro Phe Phe Ser Leu Met Ser
 145 150 155 160

Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg Pro
 165 170 175

Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn
 180 185 190


```

Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly
      195                200                205
Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly
      210                215                220
Lys Val Trp Arg Leu Glu Glu Ile Gly Gly Ser Ile Ile Gly Gly Thr
225                230                235                240
Ile Lys Ala Ile Gln Asp Lys Gly Lys Asn Pro Lys Pro Pro Arg Asp
      245                250                255
Pro Arg Leu Pro Ala Pro Lys Gly Gln Thr Val Ala Ser Phe Arg Lys
      260                265                270
Gly Leu Ala Met Leu Pro Asn Ala Ile Ala Ser Arg Leu Gly Ser Lys
      275                280                285
Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ala Asp Asn Gln
      290                295                300
Gly Tyr Val Leu Gly Tyr Glu Thr Pro Glu Gly Leu Val Ser Val Gln
305                310                315                320
Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asp Ile
      325                330                335
Leu Arg Pro Leu Ser Ile Asp Ala Ala Asp Ala Leu Ser Lys Phe Tyr
      340                345                350
Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala Ile
      355                360                365
Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln Leu
      370                375                380
His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser
385                390                395                400
Ser Leu Phe Pro Asn Arg Ala Pro Ala Gly Arg Val Leu Leu Leu Asn
      405                410                415
Tyr Ile Gly Gly Ser Thr Asn Thr Gly Ile Val Ser Lys Thr Glu Ser
      420                425                430
Asp Leu Val Gly Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile Asn
      435                440                445
Pro Arg Ala Ala Asp Pro Leu Ala Leu Gly Val Arg Val Trp Pro Gln
      450                455                460
Ala Ile Pro Gln Phe Leu Ile Gly His Leu Asp Arg Leu Ala Ala Ala
465                470                475                480
Lys Ser Ala Leu Gly Gln Gly Gly Tyr Asp Gly Leu Phe Leu Gly Gly
      485                490                495
Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala Tyr
      500                505                510
Glu Ser Ala Ser Gln Val Ser Asp Phe Leu Thr Lys Tyr Ala Tyr Lys
      515                520                525

```

<210> 11
 <211> 1847
 <212> DNA
 <213> Glycine max

<220>
 <221> CDS
 <222> (55)..(1683)
 <223> soybean protox-1

<400> 11
 ctttagcaca gtgttgaaga taacgaacga atagtgccat tactgtaacc aacc atg 57

																	Met 1	
gtt tcc gtc ttc aac gag atc cta ttc ccg ccg aac caa acc ctt ctt	105																	
Val Ser Val Phe Asn Glu Ile Leu Phe Pro Pro Asn Gln Thr Leu Leu																		
5 10 15																		
cgc ccc tcc ctc cat tcc cca acc tct ttc ttc acc tct ccc act cga	153																	
Arg Pro Ser Leu His Ser Pro Thr Ser Phe Phe Thr Ser Pro Thr Arg																		
20 25 30																		
aaa ttc cct cgc tct cgc cct aac cct att cta cgc tgc tcc att gcg	201																	
Lys Phe Pro Arg Ser Arg Pro Asn Pro Ile Leu Arg Cys Ser Ile Ala																		
35 40 45																		
gag gaa tcc acc gcg tct ccg ccc aaa acc aga gac tcc gcc ccc gtg	249																	
Glu Glu Ser Thr Ala Ser Pro Pro Lys Thr Arg Asp Ser Ala Pro Val																		
50 55 60 65																		
gac tgc gtc gtc gtc ggc gga ggc gtc agc ggc ctc tgc atc gcc cag	297																	
Asp Cys Val Val Val Gly Gly Gly Val Ser Gly Leu Cys Ile Ala Gln																		
70 75 80																		
gcc ctc gcc acc aaa cac gcc aat gcc aac gtc gtc gtc acg gag gcc	345																	
Ala Leu Ala Thr Lys His Ala Asn Ala Asn Val Val Val Thr Glu Ala																		
85 90 95																		
cga gac cgc gtc ggc ggc aac atc acc acg atg gag agg gac gga tac	393																	
Arg Asp Arg Val Gly Gly Asn Ile Thr Thr Met Glu Arg Asp Gly Tyr																		
100 105 110																		
ctc tgg gaa gaa ggc ccc aac agc ttc cag cct tct gat cca atg ctc	441																	
Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Met Leu																		
115 120 125																		
acc atg gtg gtg gac agt ggt tta aag gat gag ctt gtt ttg ggg gat	489																	
Thr Met Val Val Asp Ser Gly Leu Lys Asp Glu Leu Val Leu Gly Asp																		
130 135 140 145																		
cct gat gca cct cgg ttt gtg ttg tgg aac agg aag ttg agg ccg gtg	537																	
Pro Asp Ala Pro Arg Phe Val Leu Trp Asn Arg Lys Leu Arg Pro Val																		
150 155 160																		
ccc ggg aag ctg act gat ttg cct ttc ttt gac ttg atg agc att ggt	585																	
Pro Gly Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Gly																		
165 170 175																		
ggc aaa atc agg gct ggc ttt ggt gcg ctt gga att cgg cct cct cct	633																	
Gly Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg Pro Pro Pro																		
180 185 190																		
cca ggt cat gag gaa tcg gtt gaa gag ttt gtt cgt cgg aac ctt ggt	681																	
Pro Gly His Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly																		
195 200 205																		
gat gag gtt ttt gaa cgg ttg ata gag cct ttt tgt tca ggg gtc tat	729																	
Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr																		

210	215	220	225	
gca ggc gat cct tca aaa tta agt atg aaa gca gca ttc ggg aaa gtt				777
Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly Lys Val				
	230	235	240	
tgg aag ctg gaa aaa aat ggt ggt agc att att ggt gga act ttc aaa				825
Trp Lys Leu Glu Lys Asn Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys				
	245	250	255	
gca ata caa gag aga aat gga gct tca aaa cca cct cga gat ccg cgt				873
Ala Ile Gln Glu Arg Asn Gly Ala Ser Lys Pro Pro Arg Asp Pro Arg				
	260	265	270	
ctg cca aaa cca aaa ggt cag act gtt gga tct ttc cgg aag gga ctt				921
Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu				
	275	280	285	
acc atg ttg cct gat gca att tct gcc aga cta ggc aac aaa gta aag				969
Thr Met Leu Pro Asp Ala Ile Ser Ala Arg Leu Gly Asn Lys Val Lys				
	290	295	300	305
tta tct tgg aag ctt tca agt att agt aaa ctg gat agt gga gag tac				1017
Leu Ser Trp Lys Leu Ser Ser Ile Ser Lys Leu Asp Ser Gly Glu Tyr				
	310	315	320	
agt ttg aca tat gaa aca cca gaa gga gtg gtt tct ttg cag tgc aaa				1065
Ser Leu Thr Tyr Glu Thr Pro Glu Gly Val Val Ser Leu Gln Cys Lys				
	325	330	335	
act gtt gtc ctg acc att cct tcc tat gtt gct agt aca ttg ctg cgt				1113
Thr Val Val Leu Thr Ile Pro Ser Tyr Val Ala Ser Thr Leu Leu Arg				
	340	345	350	
cct ctg tct gct gct gct gca gat gca ctt tca aag ttt tat tac cct				1161
Pro Leu Ser Ala Ala Ala Ala Asp Ala Leu Ser Lys Phe Tyr Tyr Pro				
	355	360	365	
cca gtt gct gca gtt tcc ata tcc tat cca aaa gaa gct att aga tca				1209
Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg Ser				
	370	375	380	385
gaa tgc ttg ata gat ggt gag ttg aag ggg ttt ggt caa ttg cat cca				1257
Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro				
	390	395	400	
cgt agc caa gga gtg gaa aca tta gga act ata tac agc tca tca cta				1305
Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu				
	405	410	415	
ttc ccc aac cga gca cca cct gga agg gtt cta ctc ttg aat tac att				1353
Phe Pro Asn Arg Ala Pro Pro Gly Arg Val Leu Leu Leu Asn Tyr Ile				
	420	425	430	
gga gga gca act aat act gga att tta tcg aag acg gac agt gaa ctt				1401
Gly Gly Ala Thr Asn Thr Gly Ile Leu Ser Lys Thr Asp Ser Glu Leu				
	435	440	445	

gtg gaa aca gtt gat cga gat ttg agg aaa atc ctt ata aac cca aat 1449
 Val Glu Thr Val Asp Arg Asp Leu Arg Lys Ile Leu Ile Asn Pro Asn
 450 455 460 465

gcc cag gat cca ttt gta gtg ggg gtg aga ctg tgg cct caa gct att 1497
 Ala Gln Asp Pro Phe Val Val Gly Val Arg Leu Trp Pro Gln Ala Ile
 470 475 480

cca cag ttc tta gtt ggc cat ctt gat ctt cta gat gtt gct aaa gct 1545
 Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Asp Val Ala Lys Ala
 485 490 495

tct atc aga aat act ggg ttt gaa ggg ctc ttc ctt ggg ggt aat tat 1593
 Ser Ile Arg Asn Thr Gly Phe Glu Gly Leu Phe Leu Gly Gly Asn Tyr
 500 505 510

gtg tct ggt gtt gcc ttg gga cga tgc gtt gag gga gcc tat gag gta 1641
 Val Ser Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Val
 515 520 525

agca gct gaa gta aac gat ttt ctc aca aat aga gtg tac aaa 1683
 Ala Ala Glu Val Asn Asp Phe Leu Thr Asn Arg Val Tyr Lys
 530 535 540

tagtagcagt ttttgttttt gtggtggaat gggatgatggg actctcgtgt tccattgaat 1743

tataataatg tgaaagt ttc tcaaattcgt tcgatagggt tttggcggct tctattgctg 1803

tataatgtaaa atcctcttta agtttgaaaa aaaaaaaaaa aaaa 1847

<210> 12

<211> 543

<212> PRT

<213> Glycine max

<400> 12

Met Val Ser Val Phe Asn Glu Ile Leu Phe Pro Pro Asn Gln Thr Leu
 1 5 10 15

Leu Arg Pro Ser Leu His Ser Pro Thr Ser Phe Phe Thr Ser Pro Thr
 20 25 30

Arg Lys Phe Pro Arg Ser Arg Pro Asn Pro Ile Leu Arg Cys Ser Ile
 35 40 45

Ala Glu Glu Ser Thr Ala Ser Pro Pro Lys Thr Arg Asp Ser Ala Pro
 50 55 60

Val Asp Cys Val Val Val Gly Gly Gly Val Ser Gly Leu Cys Ile Ala
 65 70 75 80

Gln Ala Leu Ala Thr Lys His Ala Asn Ala Asn Val Val Val Thr Glu
 85 90 95

Ala Arg Asp Arg Val Gly Gly Asn Ile Thr Thr Met Glu Arg Asp Gly

100	105	110
Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Met		
115	120	125
Leu Thr Met Val Val Asp Ser Gly Leu Lys Asp Glu Leu Val Leu Gly		
130	135	140
Asp Pro Asp Ala Pro Arg Phe Val Leu Trp Asn Arg Lys Leu Arg Pro		
145	150	155
Val Pro Gly Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile		
165	170	175
Gly Gly Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg Pro Pro		
180	185	190
Pro Pro Gly His Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu		
195	200	205
Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly Val		
210	215	220
Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly Lys		
225	230	235
Val Trp Lys Leu Glu Lys Asn Gly Gly Ser Ile Ile Gly Gly Thr Phe		
245	250	255
Lys Ala Ile Gln Glu Arg Asn Gly Ala Ser Lys Pro Pro Arg Asp Pro		
260	265	270
Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys Gly		
275	280	285
Leu Thr Met Leu Pro Asp Ala Ile Ser Ala Arg Leu Gly Asn Lys Val		
290	295	300
Lys Leu Ser Trp Lys Leu Ser Ser Ile Ser Lys Leu Asp Ser Gly Glu		
305	310	315
Tyr Ser Leu Thr Tyr Glu Thr Pro Glu Gly Val Val Ser Leu Gln Cys		
325	330	335
Lys Thr Val Val Leu Thr Ile Pro Ser Tyr Val Ala Ser Thr Leu Leu		
340	345	350
Arg Pro Leu Ser Ala Ala Ala Ala Asp Ala Leu Ser Lys Phe Tyr Tyr		
355	360	365
Pro Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg		
370	375	380
Ser Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His		
385	390	395
Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser		

405	410	415
Leu Phe Pro Asn Arg Ala Pro Pro Gly Arg Val Leu Leu Leu Asn Tyr		
420	425	430
Ile Gly Gly Ala Thr Asn Thr Gly Ile Leu Ser Lys Thr Asp Ser Glu		
435	440	445
Leu Val Glu Thr Val Asp Arg Asp Leu Arg Lys Ile Leu Ile Asn Pro		
450	455	460
Asn Ala Gln Asp Pro Phe Val Val Gly Val Arg Leu Trp Pro Gln Ala		
465	470	480
Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Asp Val Ala Lys		
485	490	495
Ala Ser Ile Arg Asn Thr Gly Phe Glu Gly Leu Phe Leu Gly Gly Asn		
500	505	510
Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu		
515	520	525
Val Ala Ala Glu Val Asn Asp Phe Leu Thr Asn Arg Val Tyr Lys		
530	535	540

<210> 13
 <211> 583
 <212> DNA
 <213> Arabidopsis thaliana

<220>
 <221> misc_feature
 <222> (1) .. (583)
 <223> arabidopsis protox-1 promoter

<400> 13
 gaattccgat cgaattatat aattatcata aatttgaata agcatgttgc cttttattaa 60
 agaggtttaa taaagtttgg taataatgga ctttgacttc aaactcgatt ctcatgtaat 120
 taattaatat ttacatcaaa atttggtcac taatattacc aaattaatat actaaaatgt 180
 taattcgcaa ataaaacact aattccaaat aaagggtcat tatgataaac acgtattgaa 240
 cttgataaag caaagcaaaa ataatgggtt tcaagggttg gggttatatat gacaaaaaaa 300
 aaaaaagggt tgggttatata tctattgggc ctataaccat gttatacaaa tttgggccta 360
 actaaaataa taaaataaac gtaatggtcc tttttatatt tgggtcaaac ccaactctaa 420
 acccaaacca aagaaaaagt atacggtacg gtacacagac ttatgggtgtg tgtgattgca 480
 ggtgaatatt tctcgtcgtc ttctcctttc ttctgaagaa gattacccaa tctgaaaaaa 540
 accaagaagc tgacaaaatt ccgaattctc tgcgatttcc atg 583

<210> 14
 <211> 3848
 <212> DNA
 <213> Zea mays

<220>

<221> misc_feature
 <222> (1)..(3848)
 <223> maize protox-1 promoter

<400> 14
 tcgatctttc taggctgac cccaaatctt cctccgaagc ccctggcgcc tctgcccctt 60
 ggagctggtg gcctgaaaga gctttgctgt tgccccgaag attgtgaggt atattgtgac 120
 ctctgagact gacttccttt gtcgtcactt tgagtggagt tatggattga cctgacgtgc 180
 ctcagatgga ttcttcctcc gaagcccctg gtcatttcgg agaactctgta atcttattcc 240
 cttctttggc gaaaatctgt cagcttggat gtactcatcc atcttctgaa gcagcttctc 300
 cagagtttgt ggaggcttcc tggcgaaata ttgggctgta ggtcctggac gaagaccctt 360
 gatcatggcc tcaatgacaa tctcattggg caccgtaggc gcttgtgccc tcaatcgcaa 420
 gaaccttcgt acatatgcct gaaggtattc ttcgtgatct tgtgtgcatt ggaacagagc 480
 ctgagctgtg accgacttcg tttgaaagcc ttggaagcta gtaaccaaca tgtgcttaag 540
 cttctgccac gacgtgatag tccctggccg aagagaagaa taccatgttt gggctacatt 600
 ccgactgcc atgacgaagg acttcgccat gactacagtg ttgaccccat acgaagatat 660
 agttgcttcg tagctcatca gaaactgctt tggatctgag tgcccatcat acatggggag 720
 ctgaggtggc ttgtatgatg ggggccatgg ggtagcctgc agttctgctg ccaagggaga 780
 agcatcatca aaagtaaagg catcatgatt aaaatcatca taccatccat cctcgttgaa 840
 taagccttct tgacgaagct ccctgtgttg gggccttcga tcttgttcat cttgaacaag 900
 atgacgcact tcttcagtgg cttcgtcgat ctttcttttg agatcagcca gtcgcacat 960
 cttctccttc tttctttgta cttgttgatg gatgatctcc atgtccctga tctcttggtc 1020
 caactcctcc tcttgagtg tccagactggg ggctttcctc ttctggcttc gagcctctcg 1080
 aagagaaaga gtttcttgat ttgggtccag cggtgcagt gcagtgggtc ctggtgctga 1140
 agctttcttc ggtggcatga caaaggtcag tgcttgccga aggtgggtcga aaaggggtca 1200
 ctagaggtgg gagccaatgt tggggacttc tcaagtgcta tgagttaaga acaaggcaac 1260
 acaaaatgtt aaatattaat agctttcatc ttctgaagca ttatttccct ttgggtataa 1320
 tgatcttcag acgaaagagt cttcatcat tgcgatatat gttaatagaa ggaggagcat 1380
 atgaaatgta agagacaaca tgaacaatcg tgtagcattg ttaattcatc atcattttat 1440
 tattatggaa aaatagaaac aatattgaat tacaatgta ctttggctt gacagaagat 1500
 aaaagtacaa gcttgacgca cgagcaagta caagtcagt tgaacagtac ggggggtactg 1560
 ttcattctatt tataggcaca ggacacagcc tgtgagaaat tacagtcagt ccctttacat 1620
 ttactattga cttatagaaa aatctatgag gactggatag ctttttcccc ttttaagtcgg 1680
 tgcctttttc cgcgattaag ccgaatctcc cttgcgcata gcttcggagc atcggaacc 1740
 ttcgtcacga tcatgcccct ctcattgtgt atgcttttaa tctgaattc gaaggtacct 1800
 gtccataaac catacttgga agacattgtt aaattatgtt tttgaggacc ttcggaggac 1860
 gaaggccccc aacagtcgtg tttttgagga ccttcggaag atgaaggccc ccaacaagac 1920
 ctatccataa aaccaaccta tccacaaaac cgaccccat caccctcat ttgcctcacc 1980
 aacaacccta attaggttgt tggtttaaat tttttaggg caatttgggtc atcaccatcc 2040
 actgtcactc cacaactca atatcaataa acagactcaa tcacccaaac tgaccatacc 2100
 cataaaaccg cccaccctt ctagcgctc gccagaaacc agaaaccctg attcagagtt 2160
 caaacttaaa acgaccataa ctttcacctt ggaactcgaa tcaggtccat ttttttccaa 2220
 atcacacaaa attaaatttc gcatccgata atcaagccat ctcttacta tggttttaag 2280
 tgttgctcac actagtgtat ttatggacta atcacctgtg tatctcatac aataacatat 2340
 cagtacatct aagttgttac tcaattacca aaaccgaatt atagccttcg aaaaagggtta 2400
 tcgactagtc actcaattac caaaactaaa ctttagactt tcatgtatga catccaacat 2460
 gacactgtac tggactaaac cacccttcaa gctacacaag gagcaaaaat aactaatttt 2520
 cgtagtgtga ggagctaaag tatatgtcca caacaatagt taagggaagc cccaaggac 2580
 ttaaaagtcc ttttacctct tgaaactttt gtcgtggtct actttttcac tttaaacttc 2640
 aaaatttgac attttatcac cccttaactc ttaaaaccat ttaaattaca ttcttactag 2700
 attatagatg attttgttgt gaaaagtgtt taagacatgt ttacacattg attaaaatca 2760
 tttgttcaat ttcctagagt taaatctaatt cttattaaaa ctattagaga tactttcacg 2820
 agctctaaat atttttattt tttcattatg gaattttgtt agaattctta tagacctttt 2880
 tttgtggttt aaaagccttg ccatgttttt aacaagtttt ttttctattt tttgaaattt 2940
 tcttggaac cacttctaac ccggtagaag atttattttg ctacacttat atctacaaca 3000
 aaatcaactt atgaaattgt cttggaaact acctctaacc cggtagaatg aatttgaatg 3060
 aaaattaaac caacttacgg aatcgcccaa catatgtcga ttaaagtgga tatggataca 3120

```

tatgaagaag ccctagagat aatctaaatg gtttcagaat tgagggttat tttttgaagt 3180
ttgatgggaa gataagacca taacggtagt tcacagagat aaaaggggta tttttttcag 3240
aaatatattgt gctgcaattg atcctgtgcc tcaaattcag cctgcaacca aggccagggt 3300
ctagagcgaa caaggcccac gtcacccgtg gcccgtcagg cgaagcaggt cttgtgcaga 3360
ctttgagagg gattggatat caacggaacc aatcacgcac ggcaatgcga ttcccagccc 3420
acctgtaacg ttccagtggg ccataccttaa ctccaagccc aacggcccta ccccatctcg 3480
tcgtgtcatc cactccgccg cacaggcgct cagctccgca acgccgccgg aaatggtcgc 3540
cgccacagcc accgccatgg ccaccgctgc atcgccgcta ctcaacggga cccgaatacc 3600
tgccgcggtc cgccatcgag gactcagcgt gcgctgcgct gctgtggcgg gcggcgcggc 3660
cgaggcaccg gcatccaccg gcgcgcggct gtccgcggac tgcgttgtgg tgggcggagg 3720
catcagtggc ctctgcaccg cgcaggcgct ggccacgcgg cacggcgctc gggacgtgct 3780
tgtcacggag gcccgcgccc gccccggcgg caacattacc accgtcgagc gccccgagga 3840
agggtacc 3848

```

<210> 15

<211> 1826

<212> DNA

<213> *Gossypium hirsutum*

<220>

<221> CDS

<222> (31) .. (1647)

<223> cotton protox-1 coding sequence

<400> 15

```

cctctcgctc gcctggcccc accaccaatc atg acg gct cta atc gac ctt tct 54
Met Thr Ala Leu Ile Asp Leu Ser
1 5

ctt ctc cgt tcc atg ccc tcc gtt tcc cct ttc tcc ata ccc cac cac 102
Leu Leu Arg Ser Ser Pro Ser Val Ser Pro Phe Ser Ile Pro His His
10 15 20

cag cat ccg ccc cgc ttt cgt aaa cct ttc aag ctc cga tgc tcc ctc 150
Gln His Pro Pro Arg Phe Arg Lys Pro Phe Lys Leu Arg Cys Ser Leu
25 30 35 40

gcc gag ggt ccc acg att tcc tca tct aaa atc gac ggg gga gaa tca 198
Ala Glu Gly Pro Thr Ile Ser Ser Ser Lys Ile Asp Gly Gly Glu Ser
45 50 55

tcc atc gcg gat tgc gtc atc gtt gga ggt ggt atc agt gga ctt tgc 246
Ser Ile Ala Asp Cys Val Ile Val Gly Gly Gly Ile Ser Gly Leu Cys
60 65 70

att gct caa gct ctc gcc acc aag cac cgt gac gtc gct tcc aat gtg 294
Ile Ala Gln Ala Leu Ala Thr Lys His Arg Asp Val Ala Ser Asn Val
75 80 85

att gtg acg gag gcc aga gac cgt gtt ggt ggc aac atc act acc gtt 342
Ile Val Thr Glu Ala Arg Asp Arg Val Gly Gly Asn Ile Thr Thr Val
90 95 100

gag aga gat gga tat ctg tgg gaa gaa ggc ccc aac agt ttt cag ccc 390
Glu Arg Asp Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro
105 110 115 120

```

tcc gat cct att cta acc atg gcc gtg gat agt gga ttg aag gac gat	438
Ser Asp Pro Ile Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp	
125 130 135	
ttg gtt tta ggt gac cct aat gca ccg cga ttt gta cta tgg gag gga	486
Leu Val Leu Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly	
140 145 150	
aaa cta agg cct gtg ccc tcc aag cca acc gac ttg ccg ttt ttt gat	534
Lys Leu Arg Pro Val Pro Ser Lys Pro Thr Asp Leu Pro Phe Phe Asp	
155 160 165	
ttg atg agc att gct gga aaa ctt agg gct ggg ttc ggg gct att ggc	582
Leu Met Ser Ile Ala Gly Lys Leu Arg Ala Gly Phe Gly Ala Ile Gly	
170 175 180	
att cgg cct ccc cct ccg ggt tat gaa gaa tcg gtg gag gag ttt gtg	630
Ile Arg Pro Pro Pro Gly Tyr Glu Glu Ser Val Glu Glu Phe Val	
185 190 195 200	
cgc cgt aat ctt ggt gct gag gtt ttt gaa cgc ttt att gaa cca ttt	678
Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg Phe Ile Glu Pro Phe	
205 210 215	
tgt tca ggt gtt tat gca ggg gat cct tca aaa tta agc atg aaa gca	726
Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala	
220 225 230	
gca ttt gga aga gta tgg aag cta gaa gag att ggt ggc agc atc att	774
Ala Phe Gly Arg Val Trp Lys Leu Glu Glu Ile Gly Gly Ser Ile Ile	
235 240 245	
ggt ggc act ttc aag aca atc cag gag aga aat aag aca cct aag cca	822
Gly Gly Thr Phe Lys Thr Ile Gln Glu Arg Asn Lys Thr Pro Lys Pro	
250 255 260	
ccc aga gac ccg cgt ctg cca aaa ccg aag ggc caa aca gtt gga tct	870
Pro Arg Asp Pro Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser	
265 270 275 280	
ttt agg aag gga ctt acc atg ctg cct gag gca att gct aac agt ttg	918
Phe Arg Lys Gly Leu Thr Met Leu Pro Glu Ala Ile Ala Asn Ser Leu	
285 290 295	
ggt agc aat gta aaa tta tct tgg aag ctt tcc agt att acc aaa ttg	966
Gly Ser Asn Val Lys Leu Ser Trp Lys Leu Ser Ser Ile Thr Lys Leu	
300 305 310	
ggc aat gga ggg tat aac ttg aca ttt gaa aca cct gaa gga atg gta	1014
Gly Asn Gly Gly Tyr Asn Leu Thr Phe Glu Thr Pro Glu Gly Met Val	
315 320 325	
tct ctt cag agt aga agt gtt gta atg acc att cca tcc cat gtt gcc	1062
Ser Leu Gln Ser Arg Ser Val Val Met Thr Ile Pro Ser His Val Ala	
330 335 340	

```

agt aac ttg ttg cat cct ctc tcg gct gct gct gca gat gca tta tcc 1110
Ser Asn Leu Leu His Pro Leu Ser Ala Ala Ala Ala Asp Ala Leu Ser
345          350          355          360

caa ttt tat tat cct cca gtt gca tca gtc aca gtc tcc tat cca aaa 1158
Gln Phe Tyr Tyr Pro Pro Val Ala Ser Val Thr Val Ser Tyr Pro Lys
          365          370          375

gaa gcc att cga aaa gaa tgt ttg att gat ggt gaa ctt aag ggg ttt 1206
Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe
          380          385          390

ggc cag ttg cac cca cgc agc caa gga att gaa act tta ggg acg ata 1254
Gly Gln Leu His Pro Arg Ser Gln Gly Ile Glu Thr Leu Gly Thr Ile
          395          400          405

tac agt tca tca ctt ttc ccc aat cga gct cca tct ggc agg gtg ttg 1302
Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro Ser Gly Arg Val Leu
          410          415          420

ctc ttg aac tac ata gga gga gct acc aac act gga att ttg tcc aag 1350
Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Leu Ser Lys
425          430          435          440

act gaa ggg gaa ctt gta gaa gca gtt gat cgt gat ttg aga aaa atg 1398
Thr Glu Gly Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met
          445          450          455

ctt ata aat cct aat gca aag gat cct ctt gtt ttg ggt gta aga gta 1446
Leu Ile Asn Pro Asn Ala Lys Asp Pro Leu Val Leu Gly Val Arg Val
          460          465          470

atg gca aaa gcc att cca cag ttc ttg gtt ggt cat ttg gat ctc ctt 1494
Trp Pro Lys Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu
          475          480          485

gat agt gca aaa atg gct ctc agg gat tct ggg ttt cat gga ctg ttt 1542
Asp Ser Ala Lys Met Ala Leu Arg Asp Ser Gly Phe His Gly Leu Phe
          490          495          500

ctt ggg ggc aac tat gta tct ggt gtg gca tta gga cgg tgt gtg gaa 1590
Leu Gly Gly Asn Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Val Glu
505          510          515          520

ggt gct tac gag gtt gca gct gaa gtg aag gaa ttc ctg tca caa tat 1638
Gly Ala Tyr Glu Val Ala Ala Glu Val Lys Glu Phe Leu Ser Gln Tyr
          525          530          535

gca tac aaa taatattgaa attcttgtca ggctgcaaatt gtagaagtca 1687
Ala Tyr Lys

gttattggat agtatctctt tagctaaaaa attgggtagg gtttttttttg ttagttcctt 1747

gaccactttt tgggggttttc attagaactt catatttgta tatcatgttg caatatcaaa 1807

aaaaaaaaa aaaaaaaaaa 1826

```

<210> 16
 <211> 539
 <212> PRT
 <213> Gossypium hirsutum

<400> 16
 Met Thr Ala Leu Ile Asp Leu Ser Leu Leu Arg Ser Ser Pro Ser Val
 1 5 10 15
 Ser Pro Phe Ser Ile Pro His His Gln His Pro Pro Arg Phe Arg Lys
 20 25 30
 Pro Phe Lys Leu Arg Cys Ser Leu Ala Glu Gly Pro Thr Ile Ser Ser
 35 40 45
 Ser Lys Ile Asp Gly Gly Glu Ser Ser Ile Ala Asp Cys Val Ile Val
 50 55 60
 Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala Thr Lys
 65 70 75 80
 His Arg Asp Val Ala Ser Asn Val Ile Val Thr Glu Ala Arg Asp Arg
 85 90 95
 Val Gly Gly Asn Ile Thr Thr Val Glu Arg Asp Gly Tyr Leu Trp Glu
 100 105 110
 Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Ile Leu Thr Met Ala
 115 120 125
 Val Asp Ser Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Asn Ala
 130 135 140
 Pro Arg Phe Val Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys
 145 150 155 160
 Pro Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Ala Gly Lys Leu
 165 170 175
 Arg Ala Gly Phe Gly Ala Ile Gly Ile Arg Pro Pro Pro Pro Gly Tyr
 180 185 190
 Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly Ala Glu Val
 195 200 205
 Phe Glu Arg Phe Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp
 210 215 220
 Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly Arg Val Trp Lys Leu
 225 230 235 240
 Glu Glu Ile Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys Thr Ile Gln
 245 250 255
 Glu Arg Asn Lys Thr Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Lys
 260 265 270

Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Thr Met Leu
 275 280 285
 Pro Glu Ala Ile Ala Asn Ser Leu Gly Ser Asn Val Lys Leu Ser Trp
 290 295 300
 Lys Leu Ser Ser Ile Thr Lys Leu Gly Asn Gly Gly Tyr Asn Leu Thr
 305 310 315 320
 Phe Glu Thr Pro Glu Gly Met Val Ser Leu Gln Ser Arg Ser Val Val
 325 330 335
 Met Thr Ile Pro Ser His Val Ala Ser Asn Leu Leu His Pro Leu Ser
 340 345 350
 Ala Ala Ala Ala Asp Ala Leu Ser Gln Phe Tyr Tyr Pro Pro Val Ala
 355 360 365
 Ser Val Thr Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu
 370 375 380
 Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Ser Gln
 385 390 395 400
 Gly Ile Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn
 405 410 415
 Arg Ala Pro Ser Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala
 420 425 430
 Thr Asn Thr Gly Ile Leu Ser Lys Thr Glu Gly Glu Leu Val Glu Ala
 435 440 445
 Val Asp Arg Asp Leu Arg Lys Met Leu Ile Asn Pro Asn Ala Lys Asp
 450 455 460
 Pro Leu Val Leu Gly Val Arg Val Trp Pro Lys Ala Ile Pro Gln Phe
 465 470 475 480
 Leu Val Gly His Leu Asp Leu Leu Asp Ser Ala Lys Met Ala Leu Arg
 485 490 495
 Asp Ser Gly Phe His Gly Leu Phe Leu Gly Gly Asn Tyr Val Ser Gly
 500 505 510
 Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Val Ala Ala Glu
 515 520 525
 Val Lys Glu Phe Leu Ser Gln Tyr Ala Tyr Lys
 530 535

<210> 17
 <211> 1910
 <212> DNA

<213> Beta vulgaris

<220>

<221> CDS

<222> (1)..(1680)

<223> sugar beet protox-1 coding sequence

<400> 17

atg aaa tca atg gcg tta tca aac tgc att cca cag aca cag tgc atg	48
Met Lys Ser Met Ala Leu Ser Asn Cys Ile Pro Gln Thr Gln Cys Met	
1 5 10 15	
cca ttg cgc agc agc ggg cat tac agg ggt aat tgt atc atg ttg tca	96
Pro Leu Arg Ser Ser Gly His Tyr Arg Gly Asn Cys Ile Met Leu Ser	
20 25 30	
att cca tgt agt tta att gga aga cga ggt tat tat tca cat aag aag	144
Ile Pro Cys Ser Leu Ile Gly Arg Arg Gly Tyr Tyr Ser His Lys Lys	
35 40 45	
agg agg atg agc atg agt tgc agc aca agc tca ggc tca aag tca gcg	192
Arg Arg Met Ser Met Ser Cys Ser Thr Ser Ser Gly Ser Lys Ser Ala	
50 55 60	
gtt aaa gaa gca gga tca gga tca ggt gca gga gga ttg cta gac tgc	240
Val Lys Glu Ala Gly Ser Gly Ser Gly Ala Gly Gly Leu Leu Asp Cys	
65 70 75 80	
gta atc gtt gga ggt gga att agc ggg ctt tgc atc gcg cag gct ctt	288
Val Ile Val Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu	
85 90 95	
tgt aca aaa cac tcc tct tcc tct tta tcc cca aat ttt ata gtt aca	336
Cys Thr Lys His Ser Ser Ser Ser Leu Ser Pro Asn Phe Ile Val Thr	
100 105 110	
gag gcc aaa gac aga gtt ggc ggc aac atc gtc act gtg gag gcc gat	384
Glu Ala Lys Asp Arg Val Gly Gly Asn Ile Val Thr Val Glu Ala Asp	
115 120 125	
ggc tat atc tgg gag gag gga ccc aat agc ttc cag cct tcc gac gcg	432
Gly Tyr Ile Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Ala	
130 135 140	
gtg ctc acc atg gcg gtc gac agt ggc ttg aaa gat gag ttg gtg ctc	480
Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Glu Leu Val Leu	
145 150 155 160	
gga gat ccc aat gct cct cgc ttt gtg cta tgg aat gac aaa tta agg	528
Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Asn Asp Lys Leu Arg	
165 170 175	
ccc gta cct tcc agt ctc acc gac ctc cct ttc ttc gac ctc atg acc	576
Pro Val Pro Ser Ser Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Thr	
180 185 190	
att ccg ggc aag att agg gct gct ctt ggt gct ctc gga ttt cgc cct	624

Ile	Pro	Gly	Lys	Ile	Arg	Ala	Ala	Leu	Gly	Ala	Leu	Gly	Phe	Arg	Pro		
		195					200					205					
tct	cct	cca	cct	cat	gag	gaa	tct	gtt	gaa	cac	ttt	gtg	cgt	cgt	aat	672	
Ser	Pro	Pro	Pro	His	Glu	Glu	Ser	Val	Glu	His	Phe	Val	Arg	Arg	Asn		
		210				215					220						
ctc	gga	gat	gag	gtc	ttt	gaa	cgc	ttg	att	gaa	ccc	ttt	tgt	tca	ggt	720	
Leu	Gly	Asp	Glu	Val	Phe	Glu	Arg	Leu	Ile	Glu	Pro	Phe	Cys	Ser	Gly		
		225			230					235					240		
gtg	tat	gcc	ggt	gat	cct	gcc	aag	ctg	agt	atg	aaa	gct	gct	ttt	ggg	768	
Val	Tyr	Ala	Gly	Asp	Pro	Ala	Lys	Leu	Ser	Met	Lys	Ala	Ala	Phe	Gly		
				245				250						255			
aag	gtc	tgg	aag	ttg	gag	caa	aag	ggt	ggc	agc	ata	att	ggt	ggc	act	816	
Lys	Val	Trp	Lys	Leu	Glu	Gln	Lys	Gly	Gly	Ser	Ile	Ile	Gly	Gly	Thr		
			260					265					270				
ctc	aaa	gct	ata	cag	gaa	aga	ggg	agt	aat	cct	aag	ccg	ccc	cgt	gac	864	
Leu	Lys	Ala	Ile	Gln	Glu	Arg	Gly	Ser	Asn	Pro	Lys	Pro	Pro	Arg	Asp		
		275					280					285					
cag	cgc	ctc	cct	aaa	cca	aag	ggt	cag	act	gtt	gga	tcc	ttt	aga	aag	912	
Gln	Arg	Leu	Pro	Lys	Pro	Lys	Gly	Gln	Thr	Val	Gly	Ser	Phe	Arg	Lys		
		290				295					300						
gga	ctc	gtt	atg	ttg	cct	acc	gcc	att	tct	gct	cga	ctt	ggc	agt	aga	960	
Gly	Leu	Val	Met	Leu	Pro	Thr	Ala	Ile	Ser	Ala	Arg	Leu	Gly	Ser	Arg		
		305			310				315						320		
gtg	aaa	cta	tct	tgg	acc	ctt	tct	agt	atc	gta	aag	tca	ctc	aat	gga	1008	
Val	Lys	Leu	Ser	Trp	Thr	Leu	Ser	Ser	Ile	Val	Lys	Ser	Leu	Asn	Gly		
				325				330						335			
gaa	tat	agt	ctg	act	tat	gat	acc	cca	gat	ggc	ttg	gtt	tct	gta	aga	1056	
Glu	Tyr	Ser	Leu	Thr	Tyr	Asp	Thr	Pro	Asp	Gly	Leu	Val	Ser	Val	Arg		
			340				345					350					
acc	aaa	agt	gtt	gtg	atg	act	gtt	cca	tca	tat	gtt	gca	agt	agg	ctt	1104	
Thr	Lys	Ser	Val	Val	Met	Thr	Val	Pro	Ser	Tyr	Val	Ala	Ser	Arg	Leu		
		355					360					365					
ctt	cgt	cca	ctt	tca	gac	tct	gct	gca	gat	tct	ctt	tca	aaa	ttt	tac	1152	
Leu	Arg	Pro	Leu	Ser	Asp	Ser	Ala	Ala	Asp	Ser	Leu	Ser	Lys	Phe	Tyr		
		370				375					380						
tat	cca	cca	gtt	gca	gca	gtg	tca	ctt	tcc	tat	cct	aaa	gaa	gcg	atc	1200	
Tyr	Pro	Pro	Val	Ala	Ala	Val	Ser	Leu	Ser	Tyr	Pro	Lys	Glu	Ala	Ile		
					390					395					400		
aga	tca	gaa	tgc	ttg	att	aat	ggt	gaa	ctt	caa	ggt	ttc	ggg	caa	cta	1248	
Arg	Ser	Glu	Cys	Leu	Ile	Asn	Gly	Glu	Leu	Gln	Gly	Phe	Gly	Gln	Leu		
				405				410					415				
cat	ccc	cgc	agt	cag	ggt	gtg	gaa	acc	ttg	gga	aca	att	tat	agt	tcg	1296	
His	Pro	Arg	Ser	Gln	Gly	Val	Glu	Thr	Leu	Gly	Thr	Ile	Tyr	Ser	Ser		

420	425	430	
tct ctt ttc cct ggt cga gca cca cct ggt agg atc ttg atc ttg agc			1344
Ser Leu Phe Pro Gly Arg Ala Pro Pro Gly Arg Ile Leu Ile Leu Ser			
435	440	445	
tac atc gga ggt gct aaa aat cct ggc ata tta aac aag tcg aaa gat			1392
Tyr Ile Gly Gly Ala Lys Asn Pro Gly Ile Leu Asn Lys Ser Lys Asp			
450	455	460	
gaa ctt gcc aag aca gtt gac aag gac ctg aga aga atg ctt ata aat			1440
Glu Leu Ala Lys Thr Val Asp Lys Asp Leu Arg Arg Met Leu Ile Asn			
465	470	475	480
cct gat gca aaa ctt cct cgt gta ctg ggt gtg aga gta tgg cct caa			1488
Pro Asp Ala Lys Leu Pro Arg Val Leu Gly Val Arg Val Trp Pro Gln			
485	490	495	
gca ata ccc cag ttt tct att ggg cac ttt gat ctg ctc gat gct gca			1536
Ala Ile Pro Gln Phe Ser Ile Gly His Phe Asp Leu Leu Asp Ala Ala			
500	505	510	
aaa gct gct ctg aca gat aca ggg gtc aaa gga ctg ttt ctt ggt ggc			1584
Lys Ala Ala Leu Thr Asp Thr Gly Val Lys Gly Leu Phe Leu Gly Gly			
515	520	525	
aac tat gtt tca ggt gtt gcc ttg ggg cgg tgt ata gag ggt gct tat			1632
Asn Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala Tyr			
530	535	540	
gag tct gca gct gag gta gta gat ttc ctc tca cag tac tca gac aaa			1680
Glu Ser Ala Ala Glu Val Val Asp Phe Leu Ser Gln Tyr Ser Asp Lys			
545	550	555	560
tagagcttca gcatcctgtg taattcaaca caggcctttt tgtatctgtt gtgcgcgcat			1740
gtagtctggg cgtgggtgcta ggattgatta gttgctctgc tgtgtgatcc acaagaattt			1800
tgatggaatt tttccagatg tgggcattat atgttgctgt cttataaatc cttaatttgt			1860
acgttttagtg aattacaccg catttgatga ctaaaaaaaa aaaaaaaaaa			1910

<210> 18

<211> 560

<212> PRT

<213> Beta vulgaris

<400> 18

Met	Lys	Ser	Met	Ala	Leu	Ser	Asn	Cys	Ile	Pro	Gln	Thr	Gln	Cys	Met
1				5					10					15	

Pro	Leu	Arg	Ser	Ser	Gly	His	Tyr	Arg	Gly	Asn	Cys	Ile	Met	Leu	Ser
			20					25					30		

Ile	Pro	Cys	Ser	Leu	Ile	Gly	Arg	Arg	Gly	Tyr	Tyr	Ser	His	Lys	Lys
		35				40						45			

Arg Arg Met Ser Met Ser Cys Ser Thr Ser Ser Gly Ser Lys Ser Ala
 50 55 60
 Val Lys Glu Ala Gly Ser Gly Ser Gly Ala Gly Gly Leu Leu Asp Cys
 65 70 75 80
 Val Ile Val Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu
 85 90 95
 Cys Thr Lys His Ser Ser Ser Ser Leu Ser Pro Asn Phe Ile Val Thr
 100 105 110
 Glu Ala Lys Asp Arg Val Gly Gly Asn Ile Val Thr Val Glu Ala Asp
 115 120 125
 Gly Tyr Ile Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Ala
 130 135 140
 Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Glu Leu Val Leu
 145 150 155 160
 Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Asn Asp Lys Leu Arg
 165 170 175
 Pro Val Pro Ser Ser Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Thr
 180 185 190
 Ile Pro Gly Lys Ile Arg Ala Ala Leu Gly Ala Leu Gly Phe Arg Pro
 195 200 205
 Ser Pro Pro Pro His Glu Glu Ser Val Glu His Phe Val Arg Arg Asn
 210 215 220
 Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly
 225 230 235 240
 Val Tyr Ala Gly Asp Pro Ala Lys Leu Ser Met Lys Ala Ala Phe Gly
 245 250 255
 Lys Val Trp Lys Leu Glu Gln Lys Gly Gly Ser Ile Ile Gly Gly Thr
 260 265 270
 Leu Lys Ala Ile Gln Glu Arg Gly Ser Asn Pro Lys Pro Pro Arg Asp
 275 280 285
 Gln Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys
 290 295 300
 Gly Leu Val Met Leu Pro Thr Ala Ile Ser Ala Arg Leu Gly Ser Arg
 305 310 315 320
 Val Lys Leu Ser Trp Thr Leu Ser Ser Ile Val Lys Ser Leu Asn Gly
 325 330 335
 Glu Tyr Ser Leu Thr Tyr Asp Thr Pro Asp Gly Leu Val Ser Val Arg
 340 345 350

Thr Lys Ser Val Val Met Thr Val Pro Ser Tyr Val Ala Ser Arg Leu
 355 360 365
 Leu Arg Pro Leu Ser Asp Ser Ala Ala Asp Ser Leu Ser Lys Phe Tyr
 370 375 380
 Tyr Pro Pro Val Ala Ala Val Ser Leu Ser Tyr Pro Lys Glu Ala Ile
 385 390 395 400
 Arg Ser Glu Cys Leu Ile Asn Gly Glu Leu Gln Gly Phe Gly Gln Leu
 405 410 415
 His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser
 420 425 430
 Ser Leu Phe Pro Gly Arg Ala Pro Pro Gly Arg Ile Leu Ile Leu Ser
 435 440 445
 Tyr Ile Gly Gly Ala Lys Asn Pro Gly Ile Leu Asn Lys Ser Lys Asp
 450 455 460
 Glu Leu Ala Lys Thr Val Asp Lys Asp Leu Arg Arg Met Leu Ile Asn
 465 470 475 480
 Pro Asp Ala Lys Leu Pro Arg Val Leu Gly Val Arg Val Trp Pro Gln
 485 490 495
 Ala Ile Pro Gln Phe Ser Ile Gly His Phe Asp Leu Leu Asp Ala Ala
 500 505 510
 Lys Ala Ala Leu Thr Asp Thr Gly Val Lys Gly Leu Phe Leu Gly Gly
 515 520 525
 Asn Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala Tyr
 530 535 540
 Glu Ser Ala Ala Glu Val Val Asp Phe Leu Ser Gln Tyr Ser Asp Lys
 545 550 555 560

<210> 19

<211> 1784

<212> DNA

<213> Brassica napus

<220>

<221> CDS

<222> (47)..(1654)

<223> oilseed rape protox-1 coding sequence

<400> 19

gggccccccc caaaattgag gattctcctt ctgcgggacg atcgcc atg gat tta 55
 Met Asp Leu
 1

tct ctt ctc cgt ccg cag cca ttc cta tcg cca ttc tca aat cca ttt	103
Ser Leu Leu Arg Pro Gln Pro Phe Leu Ser Pro Phe Ser Asn Pro Phe	
5 10 15	
cct cgg tcg cgt ccc tac aag cct ctc aac ctc cgt tgc tcc gta tcc	151
Pro Arg Ser Arg Pro Tyr Lys Pro Leu Asn Leu Arg Cys Ser Val Ser	
20 25 30 35	
ggt gga tcc gtc gtc ggc tct tct aca atc gaa ggc gga gga gga ggt	199
Gly Gly Ser Val Val Gly Ser Ser Thr Ile Glu Gly Gly Gly Gly Gly	
40 45 50	
aaa acc gtc acg gcg gac tgc gtg atc gtc ggc gga gga atc agc ggc	247
Lys Thr Val Thr Ala Asp Cys Val Ile Val Gly Gly Gly Ile Ser Gly	
55 60 65	
ctg tgc att gcg caa gcg ctc gtg acg aag cac cca gac gct gca aag	295
Leu Cys Ile Ala Gln Ala Leu Val Thr Lys His Pro Asp Ala Ala Lys	
70 75 80	
aat gtg atg gtg acg gag gcg aag gac cgt gtg gga ggg aat atc atc	343
Asn Val Met Val Thr Glu Ala Lys Asp Arg Val Gly Gly Asn Ile Ile	
85 90 95	
acg cga gag gag caa ggg ttt cta tgg gaa gaa ggt ccc aat agc ttt	391
Thr Arg Glu Glu Gln Gly Phe Leu Trp Glu Glu Gly Pro Asn Ser Phe	
100 105 110 115	
cag ccg tct gat cct atg ctc act atg gtg gta gat agt ggt ttg aaa	439
Gln Pro Ser Asp Pro Met Leu Thr Met Val Val Asp Ser Gly Leu Lys	
120 125 130	
gag gat cta gtc ttg gga gat cct act gct ccg agg ttt gtg ttg tgg	487
Asp Asp Leu Val Leu Gly Asp Pro Thr Ala Pro Arg Phe Val Leu Trp	
135 140 145	
aat ggg aag ctg agg ccg gtt ccg tcg aag cta act gac ttg cct ttc	535
Asn Gly Lys Leu Arg Pro Val Pro Ser Lys Leu Thr Asp Leu Pro Phe	
150 155 160	
ttt gac ttg atg agt att gga ggg aag att aga gct ggg ttt ggt gcc	583
Phe Asp Leu Met Ser Ile Gly Gly Lys Ile Arg Ala Gly Phe Gly Ala	
165 170 175	
att ggt att cga cct tca cct ccg ggt cgt gag gaa tca gtg gaa gag	631
Ile Gly Ile Arg Pro Ser Pro Pro Gly Arg Glu Glu Ser Val Glu Glu	
180 185 190 195	
ttt gta agg cgt aat ctt ggt gat gag gtt ttt gag cgc ttg att gaa	679
Phe Val Arg Arg Asn Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu	
200 205 210	
ccc ttt tgc tca ggt gtt tat gcg gga gat cct gcg aaa ctg agt atg	727
Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ala Lys Leu Ser Met	
215 220 225	
aaa gca gct ttt ggg aag gtt tgg aag cta gag gag aat ggt ggg agc	775

Lys	Ala	Ala	Phe	Gly	Lys	Val	Trp	Lys	Leu	Glu	Glu	Asn	Gly	Gly	Ser	
	230						235					240				
atc	att	ggt	ggt	gct	ttt	aag	gca	att	caa	gcg	aaa	aat	aaa	gct	ccc	823
Ile	Ile	Gly	Gly	Ala	Phe	Lys	Ala	Ile	Gln	Ala	Lys	Asn	Lys	Ala	Pro	
	245					250					255					
aag	aca	acc	cga	gat	ccg	cgt	ctg	cca	aag	cca	aag	ggc	caa	act	gtt	871
Lys	Thr	Thr	Arg	Asp	Pro	Arg	Leu	Pro	Lys	Pro	Lys	Gly	Gln	Thr	Val	
	260				265				270						275	
ggt	tct	ttc	agg	aaa	gga	ctc	aca	atg	ctg	cca	gag	gca	atc	tcc	gca	919
Gly	Ser	Phe	Arg	Lys	Gly	Leu	Thr	Met	Leu	Pro	Glu	Ala	Ile	Ser	Ala	
				280					285					290		
agg	ttg	ggt	gac	aag	gtg	aaa	gtt	tct	tgg	aag	ctc	tca	agt	atc	act	967
Arg	Leu	Gly	Asp	Lys	Val	Lys	Val	Ser	Trp	Lys	Leu	Ser	Ser	Ile	Thr	
			295					300					305			
aag	ctg	gcc	agc	gga	gaa	tat	agc	tta	act	tac	gaa	act	ccg	gag	ggt	1015
Lys	Leu	Ala	Ser	Gly	Glu	Tyr	Ser	Leu	Thr	Tyr	Glu	Thr	Pro	Glu	Gly	
	310						315					320				
ata	gtc	act	gta	cag	agc	aaa	agt	gta	gtg	atg	act	gtg	cca	tct	cat	1063
Ile	Val	Thr	Val	Gln	Ser	Lys	Ser	Val	Val	Met	Thr	Val	Pro	Ser	His	
	325					330					335					
gtt	gct	agt	agt	ctc	ttg	cgc	cct	ctc	tct	gat	tct	gca	gct	gaa	gcg	1111
Val	Ala	Ser	Ser	Leu	Leu	Arg	Pro	Leu	Ser	Asp	Ser	Ala	Ala	Glu	Ala	
	340				345					350					355	
ctc	tca	aaa	ctc	tac	tat	ccg	cca	gtt	gca	gcc	gta	tcc	atc	tca	tac	1159
Leu	Ser	Lys	Leu	Tyr	Tyr	Pro	Pro	Val	Ala	Ala	Val	Ser	Ile	Ser	Tyr	
				360					365					370		
gcg	aaa	gaa	gca	atc	cga	agc	gaa	tgc	tta	ata	gat	ggt	gaa	cta	aaa	1207
Ala	Lys	Glu	Ala	Ile	Arg	Ser	Glu	Cys	Leu	Ile	Asp	Gly	Glu	Leu	Lys	
			375					380					385			
ggg	ttc	ggc	cag	ttg	cat	cca	cgc	acg	caa	aaa	gtg	gaa	act	ctt	gga	1255
Gly	Phe	Gly	Gln	Leu	His	Pro	Arg	Thr	Gln	Lys	Val	Glu	Thr	Leu	Gly	
	390						395					400				
aca	ata	tac	agt	tca	tcg	ctc	ttt	ccc	aac	cga	gca	ccg	cct	gga	aga	1303
Thr	Ile	Tyr	Ser	Ser	Ser	Leu	Phe	Pro	Asn	Arg	Ala	Pro	Pro	Gly	Arg	
	405					410					415					
gta	ttg	cta	ttg	aac	tac	atc	ggt	gga	gct	acc	aac	act	ggg	atc	tta	1351
Val	Leu	Leu	Leu	Asn	Tyr	Ile	Gly	Gly	Ala	Thr	Asn	Thr	Gly	Ile	Leu	
	420				425					430					435	
tca	aag	tcg	gaa	ggt	gag	tta	gtg	gaa	gca	gta	gat	aga	gac	ttg	agg	1399
Ser	Lys	Ser	Glu	Gly	Glu	Leu	Val	Glu	Ala	Val	Asp	Arg	Asp	Leu	Arg	
				440					445					450		
aag	atg	ctg	ata	aag	cca	agc	tcg	acc	gat	cca	ctt	gta	ctt	gga	gta	1447
Lys	Met	Leu	Ile	Lys	Pro	Ser	Ser	Thr	Asp	Pro	Leu	Val	Leu	Gly	Val	

455 460 465
 aaa tta tgg cct caa gcc att cct cag ttt ctg ata ggt cac att gat 1495
 Lys Leu Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly His Ile Asp
 470 475 480
 ttg gta gac gca gcg aaa gca tcg ctc tcg tca tct ggt cat gag ggc 1543
 Leu Val Asp Ala Ala Lys Ala Ser Leu Ser Ser Ser Gly His Glu Gly
 485 490 495
 tta ttc ttg ggt gga aat tac gtt gcc ggt gta gca ttg ggt cgg tgt 1591
 Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys
 500 505 510 515
 gtg gaa ggt gct tat gaa act gca acc caa gtg aat gat ttc atg tca 1639
 Val Glu Gly Ala Tyr Glu Thr Ala Thr Gln Val Asn Asp Phe Met Ser
 520 525 530
 agg tat gct tac aag taatgtaacg cagcaacgat ttgatactaa gtagtagatt 1694
 Arg Tyr Ala Tyr Lys
 535
 ttgcagtttt gactttaaga acactctgtt tgtgaaaaat tcaagtctgt gattgagtaa 1754
 atttatgtat tattactaaa aaaaaaaaaa 1784

 <210> 20
 <211> 536
 <212> SPRT
 <213> Brassica napus

 <400> 20
 Met Asp Leu Ser Leu Leu Arg Pro Gln Pro Phe Leu Ser Pro Phe Ser
 1 5 10 15
 Asn Pro Phe Pro Arg Ser Arg Pro Tyr Lys Pro Leu Asn Leu Arg Cys
 20 25 30
 Ser Val Ser Gly Gly Ser Val Val Gly Ser Ser Thr Ile Glu Gly Gly
 35 40 45
 Gly Gly Gly Lys Thr Val Thr Ala Asp Cys Val Ile Val Gly Gly Gly
 50 55 60
 Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Val Thr Lys His Pro Asp
 65 70 75 80
 Ala Ala Lys Asn Val Met Val Thr Glu Ala Lys Asp Arg Val Gly Gly
 85 90 95
 Asn Ile Ile Thr Arg Glu Glu Gln Gly Phe Leu Trp Glu Glu Gly Pro
 100 105 110
 Asn Ser Phe Gln Pro Ser Asp Pro Met Leu Thr Met Val Val Asp Ser
 115 120 125

Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Thr Ala Pro Arg Phe
 130 135 140
 Val Leu Trp Asn Gly Lys Leu Arg Pro Val Pro Ser Lys Leu Thr Asp
 145 150 155 160
 Leu Pro Phe Phe Asp Leu Met Ser Ile Gly Gly Lys Ile Arg Ala Gly
 165 170 175
 Phe Gly Ala Ile Gly Ile Arg Pro Ser Pro Pro Gly Arg Glu Glu Ser
 180 185 190
 Val Glu Glu Phe Val Arg Arg Asn Leu Gly Asp Glu Val Phe Glu Arg
 195 200 205
 Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ala Lys
 210 215 220
 Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Lys Leu Glu Glu Asn
 225 230 235 240
 Gly Gly Ser Ile Ile Gly Gly Ala Phe Lys Ala Ile Gln Ala Lys Asn
 245 250 255
 Lys Ala Pro Lys Thr Thr Arg Asp Pro Arg Leu Pro Lys Pro Lys Gly
 260 265 270
 Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Thr Met Leu Pro Glu Ala
 275 280 285
 Ile Ser Ala Arg Leu Gly Asp Lys Val Lys Val Ser Trp Lys Leu Ser
 290 295 300
 Ser Ile Thr Lys Leu Ala Ser Gly Glu Tyr Ser Leu Thr Tyr Glu Thr
 305 310 315 320
 Pro Glu Gly Ile Val Thr Val Gln Ser Lys Ser Val Val Met Thr Val
 325 330 335
 Pro Ser His Val Ala Ser Ser Leu Leu Arg Pro Leu Ser Asp Ser Ala
 340 345 350
 Ala Glu Ala Leu Ser Lys Leu Tyr Tyr Pro Pro Val Ala Ala Val Ser
 355 360 365
 Ile Ser Tyr Ala Lys Glu Ala Ile Arg Ser Glu Cys Leu Ile Asp Gly
 370 375 380
 Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Thr Gln Lys Val Glu
 385 390 395 400
 Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro
 405 410 415
 Pro Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr
 420 425 430

Gly Ile Leu Ser Lys Ser Glu Gly Glu Leu Val Glu Ala Val Asp Arg
 435 440 445
 Asp Leu Arg Lys Met Leu Ile Lys Pro Ser Ser Thr Asp Pro Leu Val
 450 455 460
 Leu Gly Val Lys Leu Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly
 465 470 475 480
 His Ile Asp Leu Val Asp Ala Ala Lys Ala Ser Leu Ser Ser Ser Gly
 485 490 495
 His Glu Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu
 500 505 510
 Gly Arg Cys Val Glu Gly Ala Tyr Glu Thr Ala Thr Gln Val Asn Asp
 515 520 525
 Phe Met Ser Arg Tyr Ala Tyr Lys
 530 535

<210> 21

<211> 1224

<212> DNA

<213> Oryza sativa

<220>

<221> CDS

<222> (1) (936)

<223> rice, protox-1 partial coding sequence

<400> 21

ccg gct ttg aag gct gca ttt ggg aag gtg tgg agg ctg gag gat act 48
 Arg Ala Leu Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Asp Thr
 1 5 10 15
 gga ggt agc att att ggt gga acc atc aag aca atc cag gag agg ggg 96
 Gly Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Gly
 20 25 30
 aaa aac ccc aaa ccg ccg agg gat ccc cgc ctt cca acg cca aag ggg 144
 Lys Asn Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Thr Pro Lys Gly
 35 40 45
 cag aca gtt gca tct ttc agg aag ggt ctg act atg ctc ccg gat gct 192
 Gln Thr Val Ala Ser Phe Arg Lys Gly Leu Thr Met Leu Pro Asp Ala
 50 55 60
 att aca tct agg ttg ggt agc aaa gtc aaa ctt tca tgg aag ttg aca 240
 Ile Thr Ser Arg Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr
 65 70 75 80
 agc att aca aag tca gac aac aaa gga tat gca tta gtg tat gaa aca 288
 Ser Ile Thr Lys Ser Asp Asn Lys Gly Tyr Ala Leu Val Tyr Glu Thr
 85 90 95

```

cca gaa ggg gtg gtc tcg gtg caa gct aaa act gtt gtc atg acc atc 336
Pro Glu Gly Val Val Ser Val Gln Ala Lys Thr Val Val Met Thr Ile
      100                      105                      110

cca tca tat gtt gct agt gat atc ttg cgg cca ctt tca agt gat gca 384
Pro Ser Tyr Val Ala Ser Asp Ile Leu Arg Pro Leu Ser Ser Asp Ala
      115                      120                      125

gca gat gct ctg tca ata ttc tat tat cca cca gtt gct gct gta act 432
Ala Asp Ala Leu Ser Ile Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr
      130                      135                      140

gtt tca tat cca aaa gaa gca att aga aaa gaa tgc tta att gac gga 480
Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly
145                      150                      155                      160

gag ctc cag ggt ttc ggc cag ctg cat ccg cgt agt cag gga gtt gag 528
Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu
      165                      170                      175

act tta gga aca ata tat agc tca tca ctc ttt cca aat cgt gct cca 576
Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro
      180                      185                      190

gct gga agg gtg tta ctt ctg aac tac ata gga ggt tct aca aat aca 624
Ala Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ser Thr Asn Thr
      195                      200                      205

ggg att gtt tcc aag act gaa agt gag ctg gta gaa gca gtt gac cgt 672
Gly Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg
      210                      215                      220

gac ctc agg aag atg ctg ata aat cct aga gca gtg gac cct ttg gtc 720
Asp Leu Arg Lys Met Leu Ile Asn Pro Arg Ala Val Asp Pro Leu Val
225                      230                      235                      240

ctt ggc gtc cgg gta tgg cca caa gcc ata cca cag ttc ctc att ggc 768
Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly
      245                      250                      255

cat ctt gat cat ctt gag gct gca aaa tct gcc ctg ggc aaa ggt ggg 816
His Leu Asp His Leu Glu Ala Ala Lys Ser Ala Leu Gly Lys Gly Gly
      260                      265                      270

tat gat gga ttg ttc ctc gga ggg aac tat gtt gca gga gtt gcc ctg 864
Tyr Asp Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu
      275                      280                      285

ggc cga tgc gtt gaa ggt gca tat gag agt gcc tca caa ata tct gac 912
Gly Arg Cys Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp
      290                      295                      300

tac ttg acc aag tac gcc tac aag tgatcaaagt tggcctgctc cttttggcac 966
Tyr Leu Thr Lys Tyr Ala Tyr Lys
305                      310

```

atagatgtga ggcttctagc agcaaaaatt tcatgggcat ctttttatcc tgattctaatt 1026
tagttagaat ttagaattgt agaggaatgt tccatttgca gttcataata gttgttcaga 1086
tttcagccat tcaatttgtg cagccattta ctatatgtag tatgatcttg taagtactac 1146
taagaacaaa tcaattatat tttcctgcaa gtgacatctt aatcgtcagc aaatccagtt 1206
actagtaaaa aaaaaaaaaa 1224

<210> 22
<211> 312
<212> PRT
<213> Oryza sativa

<400> 22

Arg Ala Leu Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Asp Thr
1 5 10 15

Gly Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Gly
20 25 30

Lys Asn Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Thr Pro Lys Gly
35 40 45

Gln Thr Val Ala Ser Phe Arg Lys Gly Leu Thr Met Leu Pro Asp Ala
50 55 60

Ile Thr Ser Arg Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr
65 70 75 80

Ser Ile Thr Lys Ser Asp Asn Lys Gly Tyr Ala Leu Val Tyr Glu Thr
85 90 95

Pro Glu Gly Val Val Ser Val Gln Ala Lys Thr Val Val Met Thr Ile
100 105 110

Pro Ser Tyr Val Ala Ser Asp Ile Leu Arg Pro Leu Ser Ser Asp Ala
115 120 125

Ala Asp Ala Leu Ser Ile Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr
130 135 140

Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly
145 150 155 160

Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu
165 170 175

Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro
180 185 190

Ala Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ser Thr Asn Thr
195 200 205

Gly Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg

210	215	220
Asp Leu Arg Lys Met Leu Ile Asn Pro Arg Ala Val Asp Pro Leu Val		
225	230	235 240
Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly		
	245	250 255
His Leu Asp His Leu Glu Ala Ala Lys Ser Ala Leu Gly Lys Gly Gly		
	260	265 270
Tyr Asp Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu		
	275	280 285
Gly Arg Cys Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp		
	290	295 300
Tyr Leu Thr Lys Tyr Ala Tyr Lys		
305	310	

<210> 23
 <211> 1590
 <212> DNA
 <213> Sorghum bicolor

<220>
 <221> CDS
 <222> (1)..(1320)
 <223> sorghum protox-1 partial coding sequence

<400> 23

tcc acc gtc gag cgc ccc gag gaa ggg tac ctc tgg gag gag ggt ccc	48
Ser Thr Val Glu Arg Pro Glu Glu Gly Tyr Leu Trp Glu Glu Gly Pro	
1 5 10 15	
aac agc ttc cag cca tcc gac ccc gtt ctc tcc atg gcc gtg gac agc	96
Asn Ser Phe Gln Pro Ser Asp Pro Val Leu Ser Met Ala Val Asp Ser	
20 25 30	
ggg ctg aag gat gac ctg gtt ttt ggg gac ccc aac gcg cca cgg ttc	144
Gly Leu Lys Asp Asp Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe	
35 40 45	
gtg ctg tgg gag ggg aag ctg agg ccc gtg cca tcc aag ccc gcc gac	192
Val Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys Pro Ala Asp	
50 55 60	
ctc ccg ttc ttc gat ctc atg agc atc cct ggc aag ctc agg gcc ggt	240
Leu Pro Phe Phe Asp Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly	
65 70 75 80	
ctc ggc gcg ctt ggc atc cgc ccg cct gct cca ggc cgc gag gag tca	288
Leu Gly Ala Leu Gly Ile Arg Pro Pro Ala Pro Gly Arg Glu Glu Ser	
85 90 95	

gtg gag gag ttt gtg cgc cgc aac ctc ggt gct gag gtc ttt gag cgc	336
Val Glu Glu Phe Val Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg	
100 105 110	
cta att gag cct ttc tgc tca ggt gtc tat gct ggc gat cct tcc aag	384
Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys	
115 120 125	
ctc agt atg aag gct gca ttt ggg aag gtg tgg cgg tta gaa gaa gct	432
Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Glu Ala	
130 135 140	
gga ggt agt att att ggt gga acc atc aag acg att cag gag agg ggc	480
Gly Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Gly	
145 150 155 160	
aag aat cca aaa cca ccg agg gat ccc cgc ctt ccg aag cca aaa ggg	528
Lys Asn Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Lys Pro Lys Gly	
165 170 175	
cag aca gtt gca tct ttc agg aag ggt ctt gcc atg ctt cca aat gcc	576
Gln Thr Val Ala Ser Phe Arg Lys Gly Leu Ala Met Leu Pro Asn Ala	
180 185 190	
atc aca tcc agc ttg ggt agt aaa gtc aaa cta tca tgg aaa ctc acg	624
Ile Thr Ser Ser Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr	
195 200 205	
agc atg aca aaa tca gat ggc aag ggg tat gtt ttg gag tat gaa aca	672
Ser Met Thr Lys Ser Asp Gly Lys Gly Tyr Val Leu Glu Tyr Glu Thr	
210 215 220	
cca gaa ggg gtt gtt ttg gtg cag gct aaa agt gtt atc atg acc att	720
Pro Glu Gly Val Val Leu Val Gln Ala Lys Ser Val Ile Met Thr Ile	
225 230 235 240	
cca tca tat gtt gct agc gac att ttg cgt cca ctt tca ggt gat gct	768
Pro Ser Tyr Val Ala Ser Asp Ile Leu Arg Pro Leu Ser Gly Asp Ala	
245 250 255	
gca gat gtt cta tca aga ttc tat tat cca cca gtt gct gct gta acg	816
Ala Asp Val Leu Ser Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr	
260 265 270	
gtt tcg tat cca aag gaa gca att aga aaa gaa tgc tta att gat ggg	864
Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly	
275 280 285	
gaa ctc cag ggt ttt ggc cag ttg cat cca cgt agt caa gga gtt gag	912
Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu	
290 295 300	
aca tta gga aca ata tac agc tca tca ctc ttt cca aat cgt gct cct	960
Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro	
305 310 315 320	
gct ggt agg gtg tta ctt cta aac tac ata gga ggt gct aca aac aca	1008

Ala Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr
 325 330 335

gga att gtt tcc aag act gaa agt gag ctg gta gaa gca gtt gac cgt 1056
 Gly Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg
 340 345 350

gac ctc cga aaa atg ctt ata aat cct aca gca gtg gac cct tta gtc 1104
 Asp Leu Arg Lys Met Leu Ile Asn Pro Thr Ala Val Asp Pro Leu Val
 355 360 365

ctt ggt gtc cga gtt tgg cca caa gcc ata cct cag ttc ctg gta gga 1152
 Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly
 370 375 380

cat ctt gat ctt ctg gag gcc gca aaa tct gcc ctg gac caa ggt ggc 1200
 His Leu Asp Leu Leu Glu Ala Ala Lys Ser Ala Leu Asp Gln Gly Gly
 385 390 395 400

tat aat ggg ctg ttc cta gga ggg aac tat gtt gca gga gtt gcc ctg 1248
 Tyr Asn Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu
 405 410 415

ggc aga tgc att gag ggc gca tat gag agt gcc gcg caa ata tat gac 1296
 Gly Arg Cys Ile Glu Gly Ala Tyr Glu Ser Ala Ala Gln Ile Tyr Asp
 420 425 430

ttc ttg acc aag tac gcc tac aag tgatggaaga agtggagcgc tgcttggttaa 1350
 Phe Leu Thr Lys Tyr Ala Tyr Lys
 435 440

ttgttatggt gcatagatga ggtgagacca ggagtagtaa aaggcgtcac gagtattttt 1410

cattcttatt ttgtaaattg cacttctggt tttttttcct gtcagtaatt agttagattt 1470

tagttatgta ggagattggt gtgttctactg ccctacaaaa gaatttttat tttgcattcg 1530

tttatgagag ctgtgcagac ttatgtaacg ttttactgta agtatcaaca aaatcaaata 1590

<210> 24

<211> 440

<212> PRT

<213> Sorghum bicolor

<400> 24

Ser Thr Val Glu Arg Pro Glu Glu Gly Tyr Leu Trp Glu Glu Gly Pro
 1 5 10 15

Asn Ser Phe Gln Pro Ser Asp Pro Val Leu Ser Met Ala Val Asp Ser
 20 25 30

Gly Leu Lys Asp Asp Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe
 35 40 45

Val Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys Pro Ala Asp
 50 55 60

Leu Pro Phe Phe Asp Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly
 65 70 75 80
 Leu Gly Ala Leu Gly Ile Arg Pro Pro Ala Pro Gly Arg Glu Glu Ser
 85 90 95
 Val Glu Glu Phe Val Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg
 100 105 110
 Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys
 115 120 125
 Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Glu Ala
 130 135 140
 Gly Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Gly
 145 150 155 160
 Lys Asn Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Lys Pro Lys Gly
 165 170 175
 Gln Thr Val Ala Ser Phe Arg Lys Gly Leu Ala Met Leu Pro Asn Ala
 180 185 190
 Ile Thr Ser Ser Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr
 195 200 205
 Ser Met Thr Lys Ser Asp Gly Lys Gly Tyr Val Leu Glu Tyr Glu Thr
 210 215 220
 Pro Glu Gly Val Val Leu Val Gln Ala Lys Ser Val Ile Met Thr Ile
 225 230 235 240
 Pro Ser Tyr Val Ala Ser Asp Ile Leu Arg Pro Leu Ser Gly Asp Ala
 245 250 255
 Ala Asp Val Leu Ser Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr
 260 265 270
 Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly
 275 280 285
 Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu
 290 295 300
 Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro
 305 310 315 320
 Ala Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr
 325 330 335
 Gly Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg
 340 345 350
 Asp Leu Arg Lys Met Leu Ile Asn Pro Thr Ala Val Asp Pro Leu Val
 355 360 365

Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly
 370 375 380
 His Leu Asp Leu Leu Glu Ala Ala Lys Ser Ala Leu Asp Gln Gly Gly
 385 390 395 400
 Tyr Asn Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu
 405 410 415
 Gly Arg Cys Ile Glu Gly Ala Tyr Glu Ser Ala Ala Gln Ile Tyr Asp
 420 425 430
 Phe Leu Thr Lys Tyr Ala Tyr Lys
 435 440

<210> 25
 <211> 93
 <212> DNA
 <213> Zea mays

<220>
 <221> misc_feature
 <222> (1)..(93)
 <223> maize protox-1 intron sequence

<400> 25
 gtacgctcct cgctggcgcc gcagcgtctt cttctcagac tcatgcgcag ccatggaatt 60
 gagatgctga atggatttta tacgcgcgcg cag 93

<210> 26
 <211> 2606
 <212> DNA
 <213> Beta vulgaris

<220>
 <221> misc_feature
 <222> (2601)..(2606)
 <223> SalI site

<220>
 <221> misc_feature
 <222> (1)..(538)
 <223> partial c-DNA of sugar beet protox-1

<220>
 <221> misc_feature
 <222> (539)..(2606)
 <223> sugar beet protox-1 promoter region (partial
 sequence of the ~3kb PstI-SalI fragment subcloned
 from pWDC-20)

<400> 26
 ctgcaggggg agggaaagag agaccgcgcac ggtgagggag gggagaccgc gacggtgagg 60

```

gaggggagaa cgcgacggtg agggagggga gaacgcgatg gtgagggagg ggagaacgcg 120
acgcgcaggg gagggggata actcgacggt gcagggagggt gagggggacg acgtgacggc 180
gcaggggagg ggggaaccgt cgcgggaagg ggaagaccgg ggggccgaca aggtggtggt 240
actggggtag ggagaggcgg cgtggagaat agtaacagag ggaggagtgg tgggtgctagg 300
gtggaagaag ggtaagaaag aggaagaaag agaattaaca ttatcttaac caaacaccac 360
tctaaatcta agggttttct tttcctttcc tctcctctcc ctttcttgat tccattccct 420
ttaccccggt gcaaccaaac gccccttat tatggaccgg aggaagtatg tagagatggt 480
cacaaaacta cttaagctgg taacttataa atatactggg tattaatga attaagtggc 540
cacaaaatga ctataaatta cttcgtaatc tttaggaact atgttgggtca cgaaataaca 600
taaaactggt tatttaaatg ctttatgtag gtactgcatt cataaatata tttctaacat 660
aatcgtggtg tgtaggtggt ttataacaca aggattaggt ttacaccaat gtcattttca 720
ttagaatgta gttagaatca ctttggaact ttgaagagtg atgacacatt tttattatgc 780
ttttatgaaa tgtctttgtg gtttttatga tagtattgag ttttaaggcaa gttggaagta 840
tatgatggag aagtacagta tataggtgac aattggtttg cttgtttcta tgagttgaaa 900
gataagtagt acacgacact gagcaatgac ctcttcttag ttgtaatttt gtcttctcga 960
cgtagtgaaa gtacaaacaa gattatggct ttcaagcttc caagataacg agattgtatg 1020
aattttgtgg tgtattttcac atcattgttt tacgttggag acaactaaa accaatgatg 1080
agtttgtgga ttcgagattt gcccctaagt cttatttacc catggcaagc atgctgaaac 1140
atgtagtca aacttacaca gctacaatgt ttagggattt tgagcaaaaa atttgggtat 1200
tctttgggta ccattatgtg agttgttgac tatggattaa acaaaatcac tatataaagt 1260
ctggaatgag aagcatccgc aattgacaca ccattgtact ttgattgttt caacaagttt 1320
attagatgta tttgtaggaa ttttgaagag gcggagatgt tgtgttataa ttgctttggg 1380
ggtgcttcac atgcactctg ttagtgagac atcttcagct tatattttta ggcggttagt 1440
gagtatgatt ttttttttcc aaacttttcc atttccatgt aattaaaaaa ggtgtttgat 1500
aaatacatgt taagatagcc aaagaaaaggc aaactttcaaa caaataaaaa aaattaagtc 1560
gcttaatcat ttttccaagt acttttttact ttttaacacca cttattactg aatctatagc 1620
cgttaagaat gcattttcac gctcacaat gcaaatcaag aacctcctca ttgaaggaga 1680
taatttagtc ctcaataacc ccgttaaaaga catttttagc atccagagaa atttcgattc 1740
agttaaaatt gcatatataa ccagagaaac aaattcagat gttagtcagt ccagctacat 1800
aggtcaatgc ctgagagt tttaaaagatcc gttatccttaa gcataagtag gtattgaggt 1860
gagttacaaa ggtaagttac ccgttacgca ccacctccac caaacaagta tggttagaag 1920
atacatgtaa tccgtttat tttagagtactat ttataaaaaa ctttttaact agaaacagtt 1980
gtttcatatt tagatataaggt taattagaat tcccgagcaa gcaagaaggg gatatagagg 2040
ataaggaggg ccgagagagc agagagagat gaaatcaatg gcgttatcaa actgcattcc 2100
acagacacag tgcattgccat tgcacagcag ccgggcattac aggggcaatt gtatcatggt 2160
gtcaattcca tgtagtttaa ttggaagacg aggttattat tcacataaga agaggaggat 2220
gagcatgagt tgcagcacia gctcaggctc aaagtcagcg gttaaagaag caggatcagg 2280
atcaggatca ggagcaggag gattgctaga ctgcgtaatc gttggaggtg gaattagcgg 2340
gctttgcac gcgcaggctc tttgtacaaa acagtcctct ttatcccaa attttatagt 2400
gacagaggcc aaagacagag ttggcggcaa catcgtcact gtggaggccg atggctatat 2460
ctgggaggag ggaccaata gcttcagcc ttccgacgcg gtgctacca tggcggtaat 2520
tctgtctct cattattcat aatcataatt caattcaatt caattcctaa cgtggaatgt 2580
ggaatgtggc atgtgcgtag gtcgac 2606

```

<210> 27

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Pclp_P1a -
plastid clpP gene promoter top strand PCR primer

<220>

<221> misc_feature

<222> (4) .. (9)

<223> EcoRI restriction site

<400> 27

gcggaattca tacttattta tcattagaaa g

31

<210> 28

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Pclp_P1b -
plastid clpP gene promoter bottom strand PCR
primer

<220>

<221> misc_feature

<222> (4)..(9)

<223> XbaI restriction site

<400> 28

gcgtctagaa agaactaaat actatatttc ac

32

<210> 29

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Pclp_P2b -
plastid clpP gene promoter bottom strand PCR
primer

<220>

<221> misc_feature

<222> (4)..(9)

<223> NcoI restriction site

<400> 29

gcgccatggt aaatgaaaga aagaactaaa

30

<210> 30

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Trps16_P1a -
plastid rps16 gene 3' untranslated region
XbaI/HindIII top strand PCR primer

<220>

<221> misc_feature

<222> (4)..(9)

<223> XbaI restriction site

<400> 30

gcgtctagat caaccgaaat tcaattaagg

30

<210> 31

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Trps16_plb -
plastid rps16 gene 3' untranslated region
XbaI/HindIII bottom strand PCR
primer

<220>

<221> misc_feature

<222> (4)..(9)

<223> HindIII restriction site

<400> 31

cgcaagcttc aatggaagca atgataa

27

<210> 32

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: minpsb_U -
plastid psbA gene 5' untranslated region 38 nt
(blunt/NcoI) including ATG start codon, top strand
primer

<400> 32

gggagtcctt gatgattaaa taaaccaaga ttttac

36

<210> 33

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: minpsb_L -
plastid psbA gene 5' untranslated region 38 nt
(blunt/NcoI) including ATG start codon (bottom
strand primer)

<400> 33

catggtaaaa tcttggttta tttaatcatc agggactccc

40

<210> 34

<211> 32
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: APRTXP1a - top strand PCR primer for amplifying the 5' portion of the mutant Arabidopsis protox gene

<220>
 <221> misc_feature
 <222> (5)..(10)
 <223> NcoI restriction site/ATG start codon

<400> 34
 gggaccatgg attgtgtgat tgtcggcgga gg 32

<210> 35
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: APRTXP1b - bottom strand PCR primer for amplifying the 5' portion of the mutant Arabidopsis protox gene

<400> 35
 ctccgctctc cagcttagtg atac 24

<210> 36
 <211> 633
 <212> DNA
 <213> sugar cane

<220>
 <221> CDS
 <222> (3)..(305)
 <223> sugar cane protox-1 partial coding sequence

<400> 36
 tt tcc aag act gaa agt gag ctg gta gaa gca gtt gac cgt gac ctc 47
 Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu
 1 5 10 15
 cgg aaa atg ctt ata aat cct aca gca gtg gac cct tta gtc ctt ggt 95
 Arg Lys Met Leu Ile Asn Pro Thr Ala Val Asp Pro Leu Val Leu Gly
 20 25 30
 gtc cga gtt tgg cca caa gcc ata cct cag ttc ctg gta gga cat ctt 143
 Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu
 35 40 45
 gat ctt ctg gag gcc gca aaa tct gcc ctg gac cga ggt ggc tac gat 191
 Asp Leu Leu Glu Ala Ala Lys Ser Ala Leu Asp Arg Gly Gly Tyr Asp

50 55 60
 ggg ctg ttc cta gga ggg aac tat gtt gca gga gtt gcc cta ggc aga 239
 Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg
 65 70 75
 tgc gtt gag ggc gcg tat gag agt gcc tcg caa ata tat gac ttc ttg 287
 Cys Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Tyr Asp Phe Leu
 80 85 90 95
 acc aag tat gcc tac aag tgatgaaaga agtggagatgc tgcttggttaa 335
 Thr Lys Tyr Ala Tyr Lys
 100
 ttgttatggt gcatagatga ggtgagacca ggagtagtaa aagcgttacg agtatttttc 395
 attcttattt tgtaaattgc acttctgggt ttttctgtc agtaattagt tagatttttag 455
 ttctgtagga gattgttctg ttcactgccc tacaaaagaa tttttatttt gcattcgttt 515
 atgagagctg tgcagactta tgtagcgttt ttctgtaagt accaacaataa tcaaatacta 575
 ttctgtaaga gctaacagaa tgtgcaactg agattgcctt ggatgaaaaa aaaaaaaaa 633
 <210> 37
 <211> 101
 <212> PRT
 <213> sugar cane
 <400> 37
 Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg
 5 10 15
 Lys Met Leu Ile Asn Pro Thr Ala Val Asp Pro Leu Val Leu Gly Val
 20 25 30
 Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp
 35 40 45
 Leu Leu Glu Ala Ala Lys Ser Ala Leu Asp Arg Gly Gly Tyr Asp Gly
 50 55 60
 Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys
 65 70 75 80
 Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Tyr Asp Phe Leu Thr
 85 90 95
 Lys Tyr Ala Tyr Lys
 100

<210> 38
 <211> 4
 <212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: sub-sequence 1

<400> 38

Ala Pro Xaa Phe

1

<210> 39

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: sub-sequence 8

<400> 39

Tyr Ile Gly Gly Xaa

1

5

<210> 40

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: sub-sequence

12

<400> 40

Ile Gly Gly Xaa

1

<210> 41

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: sub-sequence

13

<400> 41

Ser Trp Xaa Leu Xaa

1

5

<210> 42

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: sub-sequence
15

<400> 42
Gly Xaa Xaa Gly Leu
1 5

<210> 43
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: sub-sequence
17

<400> 43
Tyr Val Xaa Gly
1

<210> 44
<211> 1503
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1) .. (1503)
<223> coding sequence for a mature maize protox-1 protein

<400> 44
gct gct gtg gcg ggc ggc gcg gcc gag gca ccg gca tcc acc ggc gcg 48
Ala Ala Val Ala Gly Gly Ala Ala Glu Ala Pro Ala Ser Thr Gly Ala
1 5 10 15
cgg ctg tcc gcg gac tgc gtc gtg gtg ggc gga ggc atc agt ggc ctc 96
Arg Leu Ser Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu
20 25 30
tgc acc gcg cag gcg ctg gcc acg cgg cac ggc gtc ggg gac gtg ctt 144
Cys Thr Ala Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu
35 40 45
gtc acg gag gcc cgc gcc cgc ccc ggc ggc aac att acc acc gtc gag 192
Val Thr Glu Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr Thr Val Glu
50 55 60
cgc ccc gag gaa ggg tac ctc tgg gag gag ggt ccc aac agc ttc cag 240
Arg Pro Glu Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln
65 70 75 80
ccc tcc gac ccc gtt ctc acc atg gcc gtg gac agc gga ctg aag gat 288
Pro Ser Asp Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp
85 90 95

gac ttg gtt ttt ggg gac cca aac gcg ccg cgt ttc gtg ctg tgg gag	336
Asp Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu	
100 105 110	
ggg aag ctg agg ccc gtg cca tcc aag ccc gcc gac ctc ccg ttc ttc	384
Gly Lys Leu Arg Pro Val Pro Ser Lys Pro Ala Asp Leu Pro Phe Phe	
115 120 125	
gat ctc atg agc atc cca ggg aag ctc agg gcc ggt cta ggc gcg ctt	432
Asp Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu	
130 135 140	
ggc atc cgc ccg cct cct cca ggc cgc gaa gag tca gtg gag gag ttc	480
Gly Ile Arg Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe	
145 150 155 160	
gtg cgc cgc aac ctc ggt gct gag gtc ttt gag cgc ctc att gag cct	528
Val Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro	
165 170 175	
ttc tgc tca ggt gtc tat gct ggt gat cct tct aag ctc agc atg aag	576
Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys	
180 185 190	
gct gca ttt ggg aag gtt tgg cgg ttg gaa gaa act gga ggt agt att	624
Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Glu Thr Gly Gly Ser Ile	
195 200 205	
att ggt gga acc atc aag aca att cag gag agg agc aag aat cca aaa	672
Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Ser Lys Asn Pro Lys	
210 215 220	
cca ccg agg gat gcc cgc ctt ccg aag cca aaa ggg cag aca gtt gca	720
Pro Pro Arg Asp Ala Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Ala	
225 230 235 240	
tct ttc agg aag ggt ctt gcc atg ctt cca aat gcc att aca tcc agc	768
Ser Phe Arg Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Thr Ser Ser	
245 250 255	
ttg ggt agt aaa gtc aaa cta tca tgg aaa ctc acg agc att aca aaa	816
Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys	
260 265 270	
tca gat gac aag gga tat gtt ttg gag tat gaa acg cca gaa ggg gtt	864
Ser Asp Asp Lys Gly Tyr Val Leu Glu Tyr Glu Thr Pro Glu Gly Val	
275 280 285	
gtt tcg gtg cag gct aaa agt gtt atc atg act att cca tca tat gtt	912
Val Ser Val Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val	
290 295 300	
gct agc aac att ttg cgt cca ctt tca agc gat gct gca gat gct cta	960
Ala Ser Asn Ile Leu Arg Pro Leu Ser Ser Asp Ala Ala Asp Ala Leu	
305 310 315 320	
tca aga ttc tat tat cca ccg gtt gct gct gta act gtt tcg tat cca	1008

Ser Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro
 325 330 335
 aag gaa gca att aga aaa gaa tgc tta att gat ggg gaa ctc cag ggc 1056
 Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly
 340 345 350
 ttt ggc cag ttg cat cca cgt agt caa gga gtt gag aca tta gga aca 1104
 Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr
 355 360 365
 ata tac agt tcc tca ctc ttt cca aat cgt gct cct gac ggt agg gtg 1152
 Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp Gly Arg Val
 370 375 380
 tta ctt cta aac tac ata gga ggt gct aca aac aca gga att gtt tcc 1200
 Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser
 385 390 395 400
 aag act gaa agt gag ctg gtc gaa gca gtt gac cgt gac ctc cga aaa 1248
 Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys
 405 410 415
 atg ctt ata aat tct aca gca gtg gac cct tta gtc ctt ggt gtt cga 1296
 Met Leu Ile Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg
 420 425 430
 gtt tgg cca caa gcc ata cct cag ttc ctg gta gga cat ctt gat ctt 1344
 Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu
 435 440 445
 ctg gaa gcc gca aaa gct gcc ctg gac cga ggt ggc tac gat ggg ctg 1392
 Leu Glu Ala Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu
 450 455 460
 ttc cta gga ggg aac tat gtt gca gga gtt gcc ctg ggc aga tgc gtt 1440
 Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val
 465 470 475 480
 gag ggc gcg tat gaa agt gcc tcg caa ata tct gac ttc ttg acc aag 1488
 Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys
 485 490 495
 tat gcc tac aag tga 1503
 Tyr Ala Tyr Lys
 500

<210> 45
 <211> 500
 <212> PRT
 <213> Zea mays

<400> 45
 Ala Ala Val Ala Gly Gly Ala Ala Glu Ala Pro Ala Ser Thr Gly Ala
 1 5 10 15
 Arg Leu Ser Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu

59

Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys
 485 490 495
 Tyr Ala Tyr Lys
 500

<210> 46

<211> 1608

<212> DNA

<213> Zea mays

<220>

<221> CDS

<222> (1)..(1608)

<223> predicted coding sequence for a maize protox-1 polypeptide precursor

<400> 46

atg gtc gcc gcc aca gcc acc gcc atg gcc acc gct gca tcg ccg cta 48
 Met Val Ala Ala Thr Ala Thr Ala Met Ala Thr Ala Ala Ser Pro Leu
 1 5 10 15

ctc aac ggg acc cga ata cct gcg cgg ctc cgc cat cga gga ctc agc 96
 Leu Asn Gly Thr Arg Ile Pro Ala Arg Leu Arg His Arg Gly Leu Ser
 20 25 30

gtg cgc tgc gct gct gtg gcg ggc ggc gcg gcc gag gca ccg gca tcc 144
 Val Arg Cys Ala Ala Val Ala Gly Gly Ala Ala Glu Ala Pro Ala Ser
 35 40 45

acc ggc gcg cgg ctg tcc gcg gac tgc gtc gtg gtg ggc gga ggc atc 192
 Thr Gly Ala Arg Leu Ser Ala Asp Cys Val Val Val Gly Gly Gly Ile
 50 55 60

agt ggc ctc tgc acc gcg cag gcg ctg gcc acg cgg cac ggc gtc ggg 240
 Ser Gly Leu Cys Thr Ala Gln Ala Leu Ala Thr Arg His Gly Val Gly
 65 70 75 80

gac gtg ctt gtc acg gag gcc cgc gcc cgc ccc ggc ggc aac att acc 288
 Asp Val Leu Val Thr Glu Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr
 85 90 95

acc gtc gag cgc ccc gag gaa ggg tac ctc tgg gag gag ggt ccc aac 336
 Thr Val Glu Arg Pro Glu Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn
 100 105 110

agc ttc cag ccc tcc gac ccc gtt ctc acc atg gcc gtg gac agc gga 384
 Ser Phe Gln Pro Ser Asp Pro Val Leu Thr Met Ala Val Asp Ser Gly
 115 120 125

ctg aag gat gac ttg gtt ttt ggg gac cca aac gcg ccg cgt ttc gtg 432
 Leu Lys Asp Asp Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe Val
 130 135 140

ctg tgg gag ggg aag ctg agg ccc gtg cca tcc aag ccc gcc gac ctc 480
 Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys Pro Ala Asp Leu
 145 150 155 160

ccg ttc ttc gat ctc atg agc atc cca ggg aag ctc agg gcc ggt cta	528
Pro Phe Phe Asp Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu	
165 170 175	
ggc gcg ctt ggc atc cgc ccg cct cct cca ggc cgc gaa gag tca gtg	576
Gly Ala Leu Gly Ile Arg Pro Pro Pro Gly Arg Glu Glu Ser Val	
180 185 190	
gag gag ttc gtg cgc cgc aac ctc ggt gct gag gtc ttt gag cgc ctc	624
Glu Glu Phe Val Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg Leu	
195 200 205	
att gag cct ttc tgc tca ggt gtc tat gct ggt gat cct tct aag ctc	672
Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu	
210 215 220	
agc atg aag gct gca ttt ggg aag gtt tgg cgg ttg gaa gaa act gga	720
Ser Met Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Glu Thr Gly	
225 230 235 240	
ggt agt att att ggt gga acc atc aag aca att cag gag agg agc aag	768
Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Ser Lys	
245 250 255	
aat cca aaa cca ccg agg gat gcc cgc ctt ccg aag cca aaa ggg cag	816
Asn Pro Lys Pro Pro Arg Asp Ala Arg Leu Pro Lys Pro Lys Gly Gln	
260 265 270	
aca gtt gca tct ttc agg aag ggt ctt gcc atg ctt cca aat gcc att	864
Thr Val Ala Ser Phe Arg Lys Gly Leu Ala Met Leu Pro Asn Ala Ile	
275 280 285	
aca tcc agc ttg ggt agt aaa gtc aaa cta tca tgg aaa ctc acg agc	912
Thr Ser Ser Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr Ser	
290 295 300	
att aca aaa tca gat gac aag gga tat gtt ttg gag tat gaa acg cca	960
Ile Thr Lys Ser Asp Asp Lys Gly Tyr Val Leu Glu Tyr Glu Thr Pro	
305 310 315 320	
gaa ggg gtt gtt tcg gtg cag gct aaa agt gtt atc atg act att cca	1008
Glu Gly Val Val Ser Val Gln Ala Lys Ser Val Ile Met Thr Ile Pro	
325 330 335	
tca tat gtt gct agc aac att ttg cgt cca ctt tca agc gat gct gca	1056
Ser Tyr Val Ala Ser Asn Ile Leu Arg Pro Leu Ser Ser Asp Ala Ala	
340 345 350	
gat gct cta tca aga ttc tat tat cca ccg gtt gct gct gta act gtt	1104
Asp Ala Leu Ser Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr Val	
355 360 365	
tcg tat cca aag gaa gca att aga aaa gaa tgc tta att gat ggg gaa	1152
Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu	
370 375 380	
ctc cag ggc ttt ggc cag ttg cat cca cgt agt caa gga gtt gag aca	1200

Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu Thr
 385 390 395 400
 tta gga aca ata tac agt tcc tca ctc ttt cca aat cgt gct cct gac 1248
 Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp
 405 410 415
 ggt agg gtg tta ctt cta aac tac ata gga ggt gct aca aac aca gga 1296
 Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly
 420 425 430
 att gtt tcc aag act gaa agt gag ctg gtc gaa gca gtt gac cgt gac 1344
 Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg Asp
 435 440 445
 ctc cga aaa atg ctt ata aat tct aca gca gtg gac cct tta gtc ctt 1392
 Leu Arg Lys Met Leu Ile Asn Ser Thr Ala Val Asp Pro Leu Val Leu
 450 455 460
 ggt gtt cga gtt tgg cca caa gcc ata cct cag ttc ctg gta gga cat 1440
 Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly His
 465 470 475 480
 ctt gat ctt ctg gaa gcc gca aaa gct gcc ctg gac cga ggt ggc tac 1488
 Leu Asp Leu Leu Glu Ala Ala Lys Ala Leu Asp Arg Gly Gly Tyr
 485 490 495
 gat ggg ctg ttc cta gga ggg aac tat gtt gca gga gtt gcc ctg ggc 1536
 Asp Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly
 500 505 510
 aga tgc gtt gag ggc gcc tat gaa agt gcc tcg caa ata tct gac ttc 1584
 Arg Cys Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe
 515 520 525
 ttg acc aag tat gcc tac aag tga 1608
 Leu Thr Lys Tyr Ala Tyr Lys
 530 535

<210> 47
 <211> 535
 <212> PRT
 <213> Zea mays

<400> 47
 Met Val Ala Ala Thr Ala Thr Ala Met Ala Thr Ala Ala Ser Pro Leu
 1 5 10 15
 Leu Asn Gly Thr Arg Ile Pro Ala Arg Leu Arg His Arg Gly Leu Ser
 20 25 30
 Val Arg Cys Ala Ala Val Ala Gly Gly Ala Ala Glu Ala Pro Ala Ser
 35 40 45
 Thr Gly Ala Arg Leu Ser Ala Asp Cys Val Val Val Gly Gly Gly Ile
 50 55 60
 Ser Gly Leu Cys Thr Ala Gln Ala Leu Ala Thr Arg His Gly Val Gly
 65 70 75 80
 Asp Val Leu Val Thr Glu Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr

63